

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 20:17:15 ; Search time 268 Seconds
(without alignments)
8901.834 Million cell updates/sec

Title: US-10-829-107-3
Perfect score: 1458
Sequence: 1 gttgagagatgggatgtcc.....taataaaatgcactcattt 1458

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PGTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1458	100.0	1458	3	US-09-111-470-3
2	1458	100.0	1458	4	US-09-862-802A-3
3	1108	76.0	1370	3	US-09-111-470-9
4	1108	76.0	1370	4	US-09-862-802A-9
5	382	26.2	1224	4	US-09-949-016-4091
6	382	26.2	1277	4	US-09-016-434-1186
7	310.4	21.3	361	4	US-09-513-999C-13778
8	280.6	19.2	1348	4	US-09-949-016-4090
9	183	12.6	400	3	US-09-641-638-46
10	183	12.6	400	4	US-10-170-097-46
11	155.6	10.7	448	3	US-09-641-638-49
12	155.6	10.7	448	4	US-10-170-097-49
13	154.4	10.6	448	3	US-09-641-638-51
14	154.4	10.6	448	3	US-09-641-638-571
15	154.4	10.6	448	4	US-10-170-097-51
16	154.4	10.6	448	4	US-10-170-097-571
17	154	10.6	448	3	US-09-641-638-50
18	154	10.6	448	4	US-10-170-097-50
19	130.2	8.9	601	4	US-09-949-016-144698
20	130.2	8.9	601	4	US-09-949-016-144699
21	114.2	7.8	17723	4	US-09-949-016-15832
22	114.2	7.8	9740	4	US-09-949-016-15833
23	109.4	7.5	601	4	US-09-949-016-144713
24	76.4	5.2	601	4	US-09-949-016-144710
25	67.2	4.6	3259	5	PCT-US95-03747-1
26	61.4	4.2	2318	4	US-09-620-312D-733
27	57.4	3.9	997	4	US-09-907-794A-376

28	57.4	3.9	997	4	US-09-905-125A-376	Sequence 376, App
29	57.4	3.9	997	4	US-09-902-775A-376	Sequence 376, App
30	57.4	3.9	997	4	US-09-906-700-376	Sequence 376, App
31	57.4	3.9	997	4	US-09-903-603A-376	Sequence 376, App
32	57.4	3.9	997	4	US-09-904-920A-376	Sequence 376, App
33	57.4	3.9	997	4	US-09-909-064-376	Sequence 376, App
34	57.4	3.9	997	4	US-09-905-381A-376	Sequence 376, App
35	57.4	3.9	997	4	US-09-906-618-376	Sequence 376, App
36	53.8	3.7	417	3	US-09-535-521-10	Sequence 10, Appl
37	53.8	3.7	417	3	US-09-535-521-12	Sequence 12, Appl
38	53.8	3.7	423	3	US-09-535-521-13	Sequence 13, Appl
39	53.8	3.7	423	3	US-09-535-521-15	Sequence 15, Appl
40	53.8	3.7	561	3	US-09-535-521-16	Sequence 16, Appl
41	53.8	3.7	561	3	US-09-535-521-18	Sequence 18, Appl
42	53.8	3.7	624	3	US-09-535-521-19	Sequence 19, Appl
43	53.8	3.7	624	3	US-09-535-521-21	Sequence 21, Appl
44	53.8	3.7	876	3	US-09-535-521-4	Sequence 4, Appl
45	53.8	3.7	876	3	US-09-535-521-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-09-111-470-3
; Sequence 3, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SP0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1458 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 257..1204

```
/
/ NAME/KEY: misc_feature
/ LOCATION: 608
/ OTHER INFORMATION: /note= "short form lacks
/ OTHER INFORMATION: nucleotides 608-673"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 775
/ OTHER INFORMATION: /note= "ASGPRm (table 2) has
/ OTHER INFORMATION: sequence insert encoding GBE between nucleotides 775-776"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1064
/ OTHER INFORMATION: /note= "nucleotide 1064 of DCM2s
/ OTHER INFORMATION: may be A, which would encode asn rather than asp at the resid
/ OTHER INFORMATION: numbered 270"
/
/ US-09-111-470-3
Query Match
Best Local Similarity 100.0%; Score 1458; DB 3; Length 1458;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTGAGGAGATGGGATGTCACATGATAGGGCTCTGGGATTTTCAGACCCAGACGAC 60
DB 1 GTTGAGGAGATGGGATGTCACATGATAGGGCTCTGGGATTTTCAGACCCAGACGAC 60
QY 61 AGGACTCCAGTCACCTCTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTGAC 120
DB 61 AGGACTCCAGTCACCTCTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTGAC 120
QY 121 GTCCCACTCTGTGCTCTTGAGCAGACAAACCAAGTGGGAATTCACACCTCCAGACCTCCCA 180
DB 121 GTCCCACTCTGTGCTCTTGAGCAGACAAACCAAGTGGGAATTCACACCTCCAGACCTCCCA 180
QY 181 CAGCTCCACCCAGACTGGGGCGGCGCTGCTCCATTTTCAGTGTGACAACTTCAGT 240
DB 181 CAGCTCCACCCAGACTGGGGCGGCGCTGCTCCATTTTCAGTGTGACAACTTCAGT 240
QY 241 CCCTGTTGGCCCAAGCATGACAGGAGTATGAAAATCTCCAGTACTCTGGAGAAATAGGT 300
DB 241 CCCTGTTGGCCCAAGCATGACAGGAGTATGAAAATCTCCAGTACTCTGGAGAAATAGGT 300
QY 301 GAAAGTCCAGGGGTTTAAAAATGGCCACTTCTCTCCAGTCCCTCTCGAGCGTCTCCG 360
DB 301 GAAAGTCCAGGGGTTTAAAAATGGCCACTTCTCTCCAGTCCCTCTCGAGCGTCTCCG 360
QY 361 CTCTGGGCGCTGCATCTCTGCTGTCTGGGCGCTGGGCGCTGCTGCTGTGTCATCAT 420
DB 361 CTCTGGGCGCTGCATCTCTGCTGTCTGGGCGCTGGGCGCTGCTGCTGTGTCATCAT 420
QY 421 CTCTGTGTTGGATTCCAAAATTTCCAAAATTTTCAGAGGGACCTGCTGACCTGAGAACAGA 480
DB 421 CTCTGTGTTGGATTCCAAAATTTCCAAAATTTTCAGAGGGACCTGCTGACCTGAGAACAGA 480
QY 481 TTTTAGCACTTCACCTCAAACTCTGCGGAGATCCAGGCACTGACTTTCCAGGGCGAG 540
DB 481 TTTTAGCACTTCACCTCAAACTCTGCGGAGATCCAGGCACTGACTTTCCAGGGCGAG 540
QY 541 CAGCTTGGAAAGAACGATGATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAACG 600
DB 541 CAGCTTGGAAAGAACGATGATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAACG 600
QY 601 GCAGGCGAGGGGTATCTGAGCTCCAGGAACACTACGCAAGGACACCTTAGGCCACTG 660
DB 601 GCAGGCGAGGGGTATCTGAGCTCCAGGAACACTACGCAAGGACACCTTAGGCCACTG 660
QY 661 TCCCACTGCCCCATCTGTGTGTCTCCAGTTCATTTCTGAAATGCTCTCTGCGAGTCCAGCA 720
DB 661 TCCCACTGCCCCATCTGTGTGTCTCCAGTTCATTTCTGAAATGCTCTCTGCGAGTCCAGCA 720
QY 721 GCTGGGTGCAAGACTGAGAAACTGACCTGCCAGGTGGCTACTCTCAACAAATGCTCT 780
DB 721 GCTGGGTGCAAGACTGAGAAACTGACCTGCCAGGTGGCTACTCTCAACAAATGCTCT 780
```

```
QY 781 CACTGAAGGGACCTGCTGCCCGCTCAACTGGGTGGAGCACCAAGACAGCTGCTACTGGTT 840
DB 781 CACTGAAGGGACCTGCTGCCCGCTCAACTGGGTGGAGCACCAAGACAGCTGCTACTGGTT 840
QY 841 CTCTCACTCTGGGATGTCCTGGGCCGAGGCTGAGAAAGTACTGCGACCTGAAAGACGCCCA 900
DB 841 CTCTCACTCTGGGATGTCCTGGGCCGAGGCTGAGAAAGTACTGCGACCTGAAAGACGCCCA 900
QY 901 CTGGTGGTTCATCACTCCAGGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGC 960
DB 901 CTGGTGGTTCATCACTCCAGGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGC 960
QY 961 ATACACCTCGATGGGCTCAGTGACCTCAAGGAGCTCGAAAGTGGTGGATGGAACAGA 1020
DB 961 ATACACCTCGATGGGCTCAGTGACCTCAAGGAGCTCGAAAGTGGTGGATGGAACAGA 1020
QY 1021 CTATGCGACCGGCTTCAGAACTGGAAAGCCAGGCCAGCAGACTGGCAGGGGACGG 1080
DB 1021 CTATGCGACCGGCTTCAGAACTGGAAAGCCAGGCCAGCAGACTGGCAGGGGACGG 1080
QY 1081 GCTGGGTGAGGCGGAGCTGTGCTCACTTCCATCCAGACGGCAGGTGGAATGACGACGT 1140
DB 1081 GCTGGGTGAGGCGGAGCTGTGCTCACTTCCATCCAGACGGCAGGTGGAATGACGACGT 1140
QY 1141 CTGCCAGAGGCGCTACCACTGGGTCTGCGAGGCTGGGCTCAGACCCAGCAGGAGAG 1200
DB 1141 CTGCCAGAGGCGCTACCACTGGGTCTGCGAGGCTGGGCTCAGACCCAGCAGGAGAG 1200
QY 1201 TCAGTGAAGTCTGCTGGTGGGACCAACCGGCCACAGAAATGGCGGTGGAGGAGGACTC 1260
DB 1201 TCAGTGAAGTCTGCTGGTGGGACCAACCGGCCACAGAAATGGCGGTGGAGGAGGACTC 1260
QY 1261 TTCTCAGGACCTCTCGCAAGACCGCTCTGGGAGAGAAATAGCACTGGAGATTGGAAG 1320
DB 1261 TTCTCAGGACCTCTCGCAAGACCGCTCTGGGAGAGAAATAGCACTGGAGATTGGAAG 1320
QY 1321 CACTGTAACTATTTTGAATTTTCTTTTAAATTTTAAAGATGTTAGTGTCTTA 1380
DB 1321 CACTGTAACTATTTTGAATTTTCTTTTAAATTTTAAAGATGTTAGTGTCTTA 1380
QY 1381 AGCTTTATTTTTCCTTTTCCAACTTTTGAAGTCAACTTCATGAAGGTATAATTTTACATA 1440
DB 1381 AGCTTTATTTTTCCTTTTCCAACTTTTGAAGTCAACTTCATGAAGGTATAATTTTACATA 1440
QY 1441 ATAAAATGCACCTATT 1458
DB 1441 ATAAAATGCACCTATT 1458
```

RESULT 2

```
US-09-862-802A-3
; Sequence 3, Application US/09862802A
; Patent No. 6756478
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/09/862, 802A
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1458
```

; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid and protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (257)..(1204)
; OTHER INFORMATION: protein coding sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (608)..(673)
; OTHER INFORMATION: short form lacks nucleotides 608-673
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (775)..(776)
; OTHER INFORMATION: ASGPRM (Table 2) has sequence insert encoding GBE between nucleot
; OTHER INFORMATION: than Asp at the residue numbered 270
; OTHER INFORMATION: than Asp at the residue numbered 270
; -09-862-802A-3

Query Match 100.0%; Score 1458; DB 4; Length 1458;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTGAGGAGATGGGATGTCACAGATGATAGGGCTCTGGGATTTTCAGACCCCAAGACGAGC 60
DB 1 GTTGAGGAGATGGGATGTCACAGATGATAGGGCTCTGGGATTTTCAGACCCCAAGACGAGC 60

QY 61 AGACTCCAGTCACTTACCCAGCTCTCCAGACACAGCGCTCCAACTCTGAGTGAC 120
DB 61 AGACTCCAGTCACTTACCCAGCTCTCCAGACACAGCGCTCCAACTCTGAGTGAC 120

QY 121 GTCCACCTCTGCTCTGAGCAGACACACCTGGAATCAGACCTCCAGACCTCCCA 180
DB 121 GTCCACCTCTGCTCTGAGCAGACACACCTGGAATCAGACCTCCAGACCTCCCA 180

QY 181 CAGCTCCACCCAGACATGGGCGCGCCCTCCCTCCATTTTCAGCTGTGACAACTCAGAG 240
DB 181 CAGCTCCACCCAGACATGGGCGCGCCCTCCCTCCATTTTCAGCTGTGACAACTCAGAG 240

QY 241 CCCTGTTGGCCCAAGCATGACAGGAGTATGAAACTTCCAGTACTTGGAGAATAAGGT 300
DB 241 CCCTGTTGGCCCAAGCATGACAGGAGTATGAAACTTCCAGTACTTGGAGAATAAGGT 300

QY 301 GAAAGTCCAGGGTTTAAATGGGCCACTTCTCTCCAGTCCCTCTGAGCGTCTCCG 360
DB 301 GAAAGTCCAGGGTTTAAATGGGCCACTTCTCTCCAGTCCCTCTGAGCGTCTCCG 360

QY 361 CTCTGGGCCCTGCCATCTCTGCTGTCTGGGCGCTCGGCTGTGCTGTGGTTCATCAT 420
DB 361 CTCTGGGCCCTGCCATCTCTGCTGTCTGGGCGCTCGGCTGTGCTGTGGTTCATCAT 420

QY 421 CTGTGTGGTGGATTCCAAATTCCAAATTTTCAGAGGAGCCTGGTGACCTCCAGCAAGACA 480
DB 421 CTGTGTGGTGGATTCCAAATTCCAAATTTTCAGAGGAGCCTGGTGACCTCCAGCAAGACA 480

QY 481 TTTTAGCAACTTCACCTCAACACTGTGGCGGAGATCCAGGCACCTGATTTCCAGGGCAG 540
DB 481 TTTTAGCAACTTCACCTCAACACTGTGGCGGAGATCCAGGCACCTGATTTCCAGGGCAG 540

QY 541 CAGCTTCGAAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACG 600
DB 541 CAGCTTCGAAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACG 600

QY 601 GCAGGCGGGGTATCTGAGCTCCAGGAACACTACGAGAGGACACCTAGGCCACTG 660
DB 601 GCAGGCGGGGTATCTGAGCTCCAGGAACACTACGAGAGGACACCTAGGCCACTG 660

QY 661 TCCCACTGCCCATCTCTGTGTGTGCCAGGTTTCATTTCTGAAATGCTCTCTGCGAGTCCAGCA 720

DB 661 TCCCACTGCCCATCTCTGTGTGTGCCAGTTCATTTCTGAAATGCTCTCTGCGAGTCCAGCA 720
QY 721 GTGTGTGCAAGACCTGGAAGAACTGACCTGCGAGGTGGTACTCTCAACAACAAATGCTTC 780
DB 721 GTGTGTGCAAGACCTGGAAGAACTGACCTGCGAGGTGGTACTCTCAACAACAAATGCTTC 780
QY 781 CACTGAAGGAGCCTGCTGCCCCCTCAACTGGGTGGAGCACCAGACAGCTGCTACTGGTT 840
DB 781 CACTGAAGGAGCCTGCTGCCCCCTCAACTGGGTGGAGCACCAGACAGCTGCTACTGGTT 840
QY 841 CTCTCACTCTGGGATGCTCTGGGCCGAGCTGAGAACTACTGTCAGCTGAGAAACGCCCA 900
DB 841 CTCTCACTCTGGGATGCTCTGGGCCGAGCTGAGAACTACTGTCACGCTGAGAAACGCCCA 900
QY 901 CTTGTGGTTCATCACTCCAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGC 960
DB 901 CTTGTGGTTCATCACTCCAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGC 960
QY 961 ATACACTCGATGGGCTCAGTGACCTGAGGAGCTGGAAGTGGTGGATGGAACAGA 1020
DB 961 ATACACTCGATGGGCTCAGTGACCTGAGGAGCTGGAAGTGGTGGATGGAACAGA 1020
QY 1021 CTATGCGACCGGCTTCCAGAACTGGAAGCCAGGCCAGCAGACGACTGGCAGGGGACCG 1080
DB 1021 CTATGCGACCGGCTTCCAGAACTGGAAGCCAGGCCAGCAGACGACTGGCAGGGGACCG 1080
QY 1081 GCTGGTGGAGCGGAGGACTGTGCTCACTTCCATCCAGACGGCAGGTGGAATGACGACGT 1140
DB 1081 GCTGGTGGAGCGGAGGACTGTGCTCACTTCCATCCAGACGGCAGGTGGAATGACGACGT 1140
QY 1141 CTGCCAGAGGCCCTACCACTGGGTCTGCGAGGCTGCGCTGGTTCAGACCCAGCAGGAG 1200
DB 1141 CTGCCAGAGGCCCTACCACTGGGTCTGCGAGGCTGCGCTGGTTCAGACCCAGCAGGAG 1200
QY 1201 TCACTCAGCTGCTCTTGGTGGGACCACTGGGACCACTGGGACCACTGGGACCACTGGG 1260
DB 1201 TCACTCAGCTGCTCTTGGTGGGACCACTGGGACCACTGGGACCACTGGGACCACTGGG 1260
QY 1261 TTCTCAGACCTCTCTCGAAGACCGCTCTGGGAGAGAAATAAGCACTCTGGAGATTTGGAAG 1320
DB 1261 TTCTCAGACCTCTCTCGAAGACCGCTCTGGGAGAGAAATAAGCACTCTGGGAGATTTGGAAG 1320
QY 1321 CACTGCTAACATTTGAAATTTTCTCTTAAATTTTAAAGATGATAGTCTTTA 1380
DB 1321 CACTGCTAACATTTGAAATTTTCTCTTAAATTTTAAAGATGATAGTCTTTA 1380
QY 1381 AGCTTTTATTTTTCCTCAACTTTTCAAAAGTCACTTCAAGAGGTATATTTTACATA 1440
DB 1381 AGCTTTTATTTTTCCTCAACTTTTCAAAAGTCACTTCAAGAGGTATATTTTACATA 1440
QY 1441 ATAAAAATGCACTCATTT 1458
DB 1441 ATAAAAATGCACTCATTT 1458

RESULT 3
US-09-111-470-9
; Sequence 9, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; NUMBER OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue

ORGANISM: Human
US-09-949-016-4091

Query Match		26.2%;	Score 382;	DB 4;	Length 1224;
Best Local Similarity		66.1%;	Pred. No. 1.3e-92;		
Matches 646;		Conservative 0;	Mismatches 245;	Indels 87;	Gaps 3;
QY	224	CTGTGACAACTCAGAGCCGTGTGGGCCCAAGCATGACAGGACGATGAAACTTCCAG	283		
DB	140	CTGAGCAATCCAGGTCCAGGCCAGCCCTATCATGACCAAGGATATCAAGACCTTCAG	199		
QY	284	TACTTGAGAAATAGGTGAAAGTCCAG--GGGTTTAAAAATGGGCCACTTCCTCTCCAG	340		
DB	200	CATCTGGACAATGAGGAGAGTGACCAACCATCAGCTCAGAAAAGGGCCACTCTCTCCCCAG	259		
QY	341	TCCCTCTCGAGCGTCTCCGCTCTGGGCCCTGGCCATCTCTGCTGTCTCTGGGCTCGGC	400		
DB	260	CCCTCTCGAGCGTCTCTGCTCGGACCTCGCCTCTCTGCTCTCTCTCTGGGCTCAGC	319		
QY	401	CTGCTGCTGTGGTTCATCATCTGTGTGGTTGGATTCCAAAATTCCAAATTTACAGAGGAC	460		
DB	320	CTCTGCTGTTGTGGTTGTCTGTGTGATCGGATCCCAAACTCCAGCTGCAGGAGGAG	379		
QY	461	CTGGTGACCTTGAGAACAGATTTTAGCAACTTCACTCAAACTGTGGCGGAGATCCAG	520		
DB	380	CTCGGGGCTGAGAGAGAGCTTTCAGCAACTTTCACAGCGAGCAGGAGGCCAGGTCAAG	439		
QY	521	GCATGACTTCCAGGCGCAGGCTTGGAGAAACGATAGCATCTCTGAAAGCTGAGGTG	580		
DB	440	GGCTTGAGCACCAGGAGGCAATGTGGGAAGAAGATGAAGTCGTAGAGTCCCACTG	499		
QY	581	GAGGGTTTCAAGCAGAAACGGCAGGCGAGGGGTATCTGAGCTCCAGGAACACACTACGCAG	640		
DB	500	GAGAAACAGCAGAGGACCTGAGTG-----	524		
QY	641	AAGGCACACTAGGCCCACTGTCCCCACTGCCCATCTGTGTGTGTGCCAGTTCAATTCGAA	700		
DB	525	-----AAGATCACTCCAGC	538		
QY	701	ATGCTCTCGAGTCCAGCAGCTGGTCAAGACCTGAAAGAACTGACCTGCCAGGTGGCT	760		
DB	539	CTGCTGCTCAGCTGAAGCAGTTCTGTCTGTGACCTGCGGAGCCTGAGCTGTCAAGATGGC	598		
QY	761	ACTCTCAACAAATGCTCCACTGAAGGACCTGTGCTGCCCGTCAACTGGGTGGAGCAC	820		
DB	599	GGCTTCAGGGCAATGGCT---CAGAAAGACCTGCTGCCCGTCAACTGGGTGGAGCAC	655		
QY	821	CAAGACAGCTGTACTGGTTTCTCTCACTCTGGGATGTCTTGGGCCGAGGCTGGAAGTAC	880		
DB	656	GAGCGCAGCTGTACTGGTTTCTCTGCTCCGGGAAGGCTGGGCTGACGCCGACCAACTAC	715		
QY	881	TGCCAGCTGAAGAACGCCACCTGGTGGTTCATCACTCCAGGAGGAGCAGAAATTTGTC	940		
DB	716	TGCGGGCTGGAGGACGCGCACCTGGTGGTGGTGTCACTGCTCTGGGAGGAGCAGAAATTTGTC	775		
QY	941	CAGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGTGACCTCAAGAGGAGCCTGG	1000		
DB	776	CAGCACCATAGGCCCTGTGAACACTGGATGGGCTTCAAGCAACAAACGGGCCCTGG	835		
QY	1001	AAGTGGGTGGATGGAAACAGACTATGCGACCGGCTTCCAGAACTGGAAAGCCAGGCCACCA	1060		
DB	836	AAGTGGGTGGACGGGACGGACTACGAGACGGGCTTCAAGAACTGGAGGCCGAGCAGCCG	895		
QY	1061	GACGACTGGGAGGAGCAGGCTGGGTGGAGGGGAGGACTGTGTCATCTTCCATCCAGAC	1120		
DB	896	GACGACTGGTACGGGCCACGGGCTCGGAGGAGGCGGAGGACTGTGCCCACTTCAACGACGAC	955		
QY	1121	GGCAGGTGGAATGACAGCGTCTGCCAGAGGCCCTTACCACCTGGGTCTGCGAGGCTGGCCTG	1180		
DB	956	GGCGCTTGGAAACGACGACGCTGTGCCAGAGGCCCTTACCCTGGGTCTGCGAGACAGAGCTG	1015		
QY	1181	GGTCAGACCCAGCCAGGAG 1198			

DB1016GACAGCCAGCAGGAG1033

RESULT 6

US-09-016-434-1186

Sequence 1186, Application US/09016434

Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

TITLE OF INVENTION: PATHWAY GENE EXPRESSION

NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434

FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1186:

SEQUENCE CHARACTERISTICS:

LENGTH: 1277 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g179078

US-09-016-434-1186

Query Match26.2%; Score 382; DB 4; Length 1277;

Best Local Similarity66.1%; Pred. No. 1.3e-92;

Matches 646; Conservative 0; Mismatches 245; Indels 87;

QY224CTGTGACAACTCAGAGCCGCTGTGGCCCAAGCATGACAGGACGCTATGAAAC

DB140CTGAGCAATCCAGGTCAGCGCCATCATGACCAAGAGTATCAGACCTTCAG

QY284TACTTGGAGAATAAGGTGAAGTCCAG---GGGTTTAAAAATGGGCCACTTCTCTCCAG

DB200CATCTGGACAATGAGGAGAGTACCACCATCAGCTCAGAAAGGGCCACTCTCTCCAG

QY341TCCCTCTCGAGGCTCTCGCTCTGGCCCTGCGATCTCTGCTGCTCCCTGGCCCTCGC

DB260CCCTCTCTGAGGCTCTGCTCTCGGACCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT

QY401CTGCTGCTGCTCATCTCTGCTGCTGGATTCCAAATTTCCAGAGGAC

DB320CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

QY461CTGGTGAACCTGAGAACAGATTTTAGCACTTCACTCAAACTGTGGCGGAGATCCAG


```

; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 46
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 128
; OTHER INFORMATION: 12-198-128 : polymorphic base A or G
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 108..127
; OTHER INFORMATION: 12-198-128.misl, potential
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1..20
; OTHER INFORMATION: upstream amplification primer
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 380..400
; OTHER INFORMATION: downstream amplification primer, complement
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 116..140
; OTHER INFORMATION: 12-198-128 potential probe
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 249
; OTHER INFORMATION: n=a, g, c or t
US-10-170-097-46

Query Match      12.6%; Score 183; DB 4; Length 400;
Best Local Similarity 98.9%; Pred. No. 3.le-39;
Matches 183; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GTTGAGGAGATGGATGTCCTCCAGATGATAGGGCTCTCTGGGATTTTCAGACCCCAAGACCAGC 60
Db      8 GTTGAGGAGATGGATGTCCTCCAGATGATAGGGCTCTCTGGGATTTTCAGACCCCAAGACCAGC 67

Qy     61 AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC 120
Db     68 AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC 127

Qy    121 GTCCCACTCTGTCCTTGAGCACAACCAACGTGGGAATCACACCTTCAGACCTCCCA 180
Db    128 RTCCCACTCTGTCCTTGAGCACAACCAACGTGGGAATCACACCTTCAGACCTCCCA 187

Qy    181 CAGCT 185
Db    188 CAGT 192

```

RESULT 11

```

US-09-641-638-49
; Sequence 49, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.05ICP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 49
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 129
; OTHER INFORMATION: 12-214-129 : polymorphic base C or T
; NAME/KEY: misc_binding
; LOCATION: 110..128
; OTHER INFORMATION: 12-214-129.misl
; NAME/KEY: misc_binding
; LOCATION: 130..149
; OTHER INFORMATION: 12-214-129.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 1..20
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 429..448
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 117..141
; OTHER INFORMATION: 12-214-129 potential probe
; OTHER INFORMATION: 12-214-129 potential probe
US-09-641-638-49

```

```

Query Match      10.7%; Score 155.6; DB 3; Length 448;
Best Local Similarity 99.4%; Pred. No. 8e-32;
Matches 155; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      776 GCTCCACTGAAGGACCTCTGCCCGTCAACTGGGTGAGCACCAGACAGCTGCTAC 835
Db     103 GCCTCCACTGAAGGACCTCTGCCCYGTCAACTGGGTGAGCACCAGACAGCTGCTAC 162

Qy     836 TGCTTCTCTCACTCTGGGATGTCCTGGCCGAGGCTGAGAAGTACTGCCAGCTGAAGAC 895
Db     163 TGCTTCTCTCACTCTGGGATGTCCTGGCCGAGGCTGAGAAGTACTGCCAGCTGAAGAC 222

Qy     896 GCCCACCCTGGTGGTTCATCAACTCCAGGGAGGAGCAG 931
Db     223 GCCCACCCTGGTGGTTCATCAACTCCAGGGAGGAGCAG 258

```

RESULT 12

```

US-10-170-097-49
; Sequence 49, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:

```

APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GEN-T114XC2D1
CURRENT APPLICATION NUMBER: US/10/170,097
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 49
LENGTH: 448
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 129
OTHER INFORMATION: 12-214-129 : polymorphic base C or T
FEATURE:
NAME/KEY: misc_binding
LOCATION: 110..128
OTHER INFORMATION: 12-214-129.mis1
FEATURE:
NAME/KEY: misc_binding
LOCATION: 130..149
OTHER INFORMATION: 12-214-129.mis2, potential complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1..20
OTHER INFORMATION: upstream amplification primer
FEATURE:
NAME/KEY: primer_bind
LOCATION: 429..448
OTHER INFORMATION: downstream amplification primer, complement
FEATURE:
NAME/KEY: misc_binding
LOCATION: 117..141
OTHER INFORMATION: 12-214-129 potential probe
US-10-170-097-49

Query Match 10.7%; Score 155.6; DB 4; Length 448;
Best Local Similarity 99.4%; Pred. No. 8e-32; Indels 0; Gaps 0;
Matches 155; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 776 GCCTCCACTGAAGGACCTGCTGCCCGTCAACTGGGTGGAGCACCAGACAGCTGCTAC 835
DB 103 GCCTCCACTGAAGGACCTGCTGCCCGTCAACTGGGTGGAGCACCAGACAGCTGCTAC 162

QY 836 TGGTTCTCTCACTCTGGATGTCCTGGGCCGAGGCTGAGAAGTACTGCCAGCTGAAGAAC 895
DB 163 TGGTTCTCTCACTCTGGATGTCCTGGGCCGAGGCTGAGAAGTACTGCCAGCTGAAGAAC 222

QY 896 GCCCACTGTGTGTCATCACTCCAGGAGGAGCAG 931
DB 223 GCCCACTGTGTGTCATCACTCCAGGAGGAGCAG 258

RESULT 13
US-09-641-638-51
Sequence 51, Application US/09641638
Patent No. 6432648
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GENSET.051CP1

APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GENSET.051CP1
CURRENT APPLICATION NUMBER: US/09/641,638
CURRENT FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 51
LENGTH: 448
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 358
OTHER INFORMATION: 12-214-360 : polymorphic base G or C
NAME/KEY: misc_binding
LOCATION: 338..357
OTHER INFORMATION: 12-214-360.mis1, potential
NAME/KEY: misc_binding
LOCATION: 359..378
OTHER INFORMATION: 12-214-360.mis2, potential complement
NAME/KEY: primer_bind
LOCATION: 1..20
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer_bind
LOCATION: 429..448
OTHER INFORMATION: downstream amplification primer, complement
NAME/KEY: misc_binding
LOCATION: 346..370
OTHER INFORMATION: 12-214-360 potential probe
US-09-641-638-51

Query Match 10.6%; Score 154.4; DB 3; Length 448;
Best Local Similarity 99.4%; Pred. No. 1.7e-31;
Matches 155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 776 GCCTCCACTGAAGGACCTGCTGCCCGTCAACTGGGTGGAGCACCAGACAGCTGCTAC 835
DB 103 GCCTCCACTGAAGGACCTGCTGCCCGTCAACTGGGTGGAGCACCAGACAGCTGCTAC 162

QY 836 TGGTTCTCTCACTCTGGATGTCCTGGGCCGAGGCTGAGAAGTACTGCCAGCTGAAGAAC 895
DB 163 TGGTTCTCTCACTCTGGATGTCCTGGGCCGAGGCTGAGAAGTACTGCCAGCTGAAGAAC 222

QY 896 GCCCACTGTGTGTCATCACTCCAGGAGGAGCAG 931
DB 223 GCCCACTGTGTGTCATCACTCCAGGAGGAGCAG 258

RESULT 14
US-09-641-638-571
Sequence 571, Application US/09641638
Patent No. 6432648
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GENSET.051CP1

CURRENT APPLICATION NUMBER: US/09/641,638
CURRENT FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 571
LENGTH: 448
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 85
OTHER INFORMATION: 12-214-85 : deletion CCTAT
NAME/KEY: misc_binding
LOCATION: 65..84
OTHER INFORMATION: 12-214-85.misl, potential
NAME/KEY: primer_bind
LOCATION: 1..20
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer_bind
LOCATION: 429..448
OTHER INFORMATION: downstream amplification primer, complement
US-09-641-638-571

Query Match 10.6%; Score 154.4; DB 3; Length 448;
Best Local Similarity 99.4%; Pred. No. 1.7e-31;
Matches 155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 776 GCCTCCACTGAAGGGACCTGCTGCCCGTCAACTGGGTGAGCACCACCAAGACAGCTGCTAC 835
Db 103 GCCTCCACTGAAGGGACCTGCTGCCCTGTCACTGGGTGAGCACCACCAAGACAGCTGCTAC 162
Qy 836 TGGTTCTCTCACTCTGGGATGTCTGGCGGAGGCTGAGAAGTACTGCCACAGCTGAAGAAC 895
Db 163 TGGTTCTCTCACTCTGGGATGTCTGGCGGAGGCTGAGAAGTACTGCCACAGCTGAAGAAC 222
Qy 896 GCCCACCCTGGTGTCTATCACTCAACTCCAGGGAGGAGCAG 931
Db 223 GCCCACCCTGGTGTCTATCACTCAACTCCAGGGAGGAGCAG 258

RESULT 15
US-10-170-097-51
Sequence 51, Application US/10170097
Patent No. 6794143
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GEN-T114XC2D1
CURRENT APPLICATION NUMBER: US/10/170,097
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304

SOFTWARE: Patent.pm
SEQ ID NO 51
LENGTH: 448
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 358
OTHER INFORMATION: 12-214-360 : polymorphic base G or C
FEATURE:
NAME/KEY: misc_binding
LOCATION: 338..357
OTHER INFORMATION: 12-214-360.misl, potential
FEATURE:
NAME/KEY: misc_binding
LOCATION: 359..378
OTHER INFORMATION: 12-214-360.misl2, potential complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1..20
OTHER INFORMATION: upstream amplification primer
FEATURE:
NAME/KEY: primer_bind
LOCATION: 429..448
OTHER INFORMATION: downstream amplification primer, complement
FEATURE:
NAME/KEY: misc_binding
LOCATION: 346..370
OTHER INFORMATION: 12-214-360 potential probe
US-10-170-097-51

Query Match 10.6%; Score 154.4; DB 4; Length 448;
Best Local Similarity 99.4%; Pred. No. 1.7e-31;
Matches 155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 776 GCCTCCACTGAAGGGACCTGCTGCCCGTCAACTGGGTGAGCACCACCAAGACAGCTGCTAC 835
Db 103 GCCTCCACTGAAGGGACCTGCTGCCCTGTCACTGGGTGAGCACCACCAAGACAGCTGCTAC 162
Qy 836 TGGTTCTCTCACTCTGGGATGTCTGGCGGAGGCTGAGAAGTACTGCCACAGCTGAAGAAC 895
Db 163 TGGTTCTCTCACTCTGGGATGTCTGGCGGAGGCTGAGAAGTACTGCCACAGCTGAAGAAC 222
Qy 896 GCCCACCCTGGTGTCTATCACTCAACTCCAGGGAGGAGCAG 931
Db 223 GCCCACCCTGGTGTCTATCACTCAACTCCAGGGAGGAGCAG 258

Search completed: May 27, 2005, 23:57:16
Job time : 271 secs

This Page Blank (uspto)

Result No.	Query	Score	Match Length			DB	ID	Description
			Match	Length	%			
C	2	1277.2	87.6	1610	3	CR620226	full-leng	CR620226 full-leng
	3	841.6	57.7	1134	5	BQ067335	AGENCY	BQ067335 AGENCY
	4	759.4	52.1	816	5	BQ018186	UI-H-DRI	BQ018186 UI-H-DRI
	5	736	50.5	827	1	ALS46075	ALS46075	ALS46075 ALS46075
	6	709.6	48.7	842	4	BI920020	603035394	BI920020 603035394
C	7	706.2	48.4	829	4	BI821633	603036669	BI821633 603036669
	8	658.4	45.2	678	5	BUG21880	UI-H-DF0	BUG21880 UI-H-DF0
C	8	656	45	674	6	CA412958	UI-H-E20	CA412958 UI-H-E20
	9	655.8	45.0	721	4	BI938214	603083132	BI938214 603083132
C	10	654.8	44.9	710	5	BUG687211	UI-CF-EC1	BUG687211 UI-CF-EC1
	11	654	44.9	774	4	BI920664	603034446	BI920664 603034446
C	12	649	44.5	728	4	BI838988	603087264	BI838988 603087264
	13	637	43.7	663	6	CA418829	UI-H-E21	CA418829 UI-H-E21
C	14	621.4	42.6	634	4	BI905917	603063080	BI905917 603063080
	15	616.6	42.3	816	4	BI767532	603061425	BI767532 603061425
C	16	604.6	41.5	852	4	BI910159	603069529	BI910159 603069529
	17	604	41.4	671	4	BM714409	UI-E-EJ0	BM714409 UI-E-EJ0
C	18	592.6	40.6	935	4	BI912448	603291140	BI912448 603291140
	19	592.2	40.6	720	4	BI910082	603067920	BI910082 603067920
C	20	582.4	39.9	824	4	BI838052	603086617	BI838052 603086617
	21	578	39.6	721	4	BI518752	603061934	BI518752 603061934
C	22	567.4	38.9	718	4	BI910338	603068394	BI910338 603068394
	23	565.8	38.8	583	5	BP228397	BP228397	BP228397 BP228397
C	24	563.8	38.7	600	5	BUG16457	UI-H-DF0	BUG16457 UI-H-DF0

ב"ר

ch 87.6%; Score 1277.2; DB 3; Length 1610;
 1 similarity 94.2%; Pred. No. 0;
 3/50; Conservative 0; Mismatches 3; Indels 81; Gaps 1;
 1 AGGAGATGGGGATGTCCAGATGATAGGGTCTCGGGATTTGAGACCCAGACCGAGGGA 64
 1 AGGAGATGGGGATGTCCAGATGATAGGGTCTCGGGATTTGAGACCCAGACCGAGGGA 60

QY 65 CTCAGTCACTTACCCAGCTCTCCAGGACACAGCGTCCCAAATCTCTGAGTGAGCTCC 124
DB 61 CTCAGTCACTTACCCAGCTCTCCAGGACACAGCGTCCCAAATCTCTGAGTGAGCTCC 120
QY 125 CAGCTCTGGTCTTGAGGACAAACAAAGTGGGAATCACACCTCCAGACCTCCCAACAGC 184
DB 121 CAGCTCTGGTCTTGAGGACAAACAAAGTGGGAATCACACCTCCAGACCTCCCAACAGC 180
QY 185 TCCACCCAGACTGGGCGCGGCGCTCCCTCCATTTTCAGTGTGACAACTCAGAGCGCT 244
DB 181 TCCACCCAGACTGGGCGCGGCGCTCCCTCCATTTTCAGTGTGACAACTCAGAGCGCT 240
QY 245 GTTGGGCCAAGCATGACAGGACGTATGAAATCTCCAGTACTTGGAGAAATAGGTGAAA 304
DB 241 GTTGGGCCAAGCATGACAGGACGTATGAAATCTCCAGTACTTGGAGAAATAGGTGAAA 300
QY 305 GTCAGGGGTTAAAAATAGGCGCACTTCCTCTCCAGTCCCTCTGAGCGGTCTCCGCTCT 364
DB 301 GTCAGGGGTTAAAAATAGGCGCACTTCCTCTCCAGTCCCTCTGAGCGGTCTCTGCTCT 360
QY 365 GGGCCCTGCACTCTCTGCTGCTCCCTGGGCGCTCGGCTGCTGCTGCTGCTCATCTCT 424
DB 361 GGGCCCTGCACTCTCTGCTGCTCCCTGGGCGCTCGGCTGCTGCTGCTGCTCATCTCT 420
QY 425 GTGGTTGGATTTCCAAATTTCCAGAGGAGCTGTGTGACCTCTGAGAACAGATTTT 484
DB 421 GTGGTTGGATTTCCAAATTTCCAGAGGAGCTGTGTGACCTCTGAGAACAGATTTT 480
QY 485 AGCAACTTCACTCAACACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGCAGCAGC 544
DB 481 AGCAACTTCACTTCAACACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGCAGCAGC 540
QY 545 TTGGAGAACAGTACATCTCTGAAGCTGAGGTGAGGTTTCAAGCAGGAACGGCAG 604
DB 541 TTGGAGAACAGTACATCTCTGAAGCTGAGGTGAGGTTTCAAGCAGGAACGGCAG 600
QY 605 GCAGGGGTATCTGAGCTCCAGGAACACACTACGACAGAGCACACCTAGGCCACTGTCC 664
DB 601 G----- 601
QY 665 CACTGCCCATCTGTGTGTCTGAGTTCATTTGAAATGCTCTGCGAGTCCAGCAGCTG 724
DB 602 -----CAGTTTCATTTCTGAAATGCTCTGCGAGTCCAGCAGCTG 639
QY 725 GTGCAAGACTGAAGAACTGACCTGCCAGGTGGCTACTCTCAACCAATGCTCCACT 784
DB 640 GTGCAAGACTGAAGAACTGACCTGCCAGGTGGCTACTCTCAACCAATGCTCCACT 699
QY 785 GAAGGACCTGTGTCGCCCGTCAACTGGGTGGAGCACCAAGCAGCTGTCTGTTCTCT 844
DB 700 GAAGGACCTGTGTCGCCCGTCAACTGGGTGGAGCACCAAGCAGCTGTCTGTTCTCT 759
QY 845 CACTCTGGAGTGTCTGGGCGAGGTGAGAACTATGTCGAGCTGAAGAACGCCACCTG 904
DB 760 CACTCTGGAGTGTCTGGGCGAGGTGAGAACTATGTCGAGCTGAAGAACGCCACCTG 819
QY 905 GTGGTTCATCAACTCCAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGATAC 964
DB 820 GTGGTTCATCAACTCCAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGATAC 879
QY 965 ACCTGGATGGGCTCTAGTCACTCCAGGAGCTGGAAGTGGGTGGATGGAAACAGACTAT 1024
DB 880 ACCTGGATGGGCTCTAGTCACTCCAGGAGCTGGAAGTGGGTGGATGGAAACAGACTAT 939
QY 1025 GCGACCGGCTTCAGAACTGGAAAGCCAGGCGACCGACAGACTGTGGGAGGCGAGGGCTG 1084
DB 940 GCGACCGGCTTCAGAACTGGAAAGCCAGGCGACCGACAGACTGTGGGAGGCGAGGGCTG 999
QY 1085 GGTGGAGGCGAGGACTGTGCTCACTTCCATCCAGAGCGGAGGTGGAAATGACGAGCTGTC 1144
DB 1000 GGTGGAGGCGAGGACTGTGCTCACTTCCATCCAGAGCGGAGGTGGAAATGACGAGCTGTC 1059

QY 1145 CAGAGCCCTACCCTGGGTCTGCGAGGCTGGCTGGGTGAGAGGAGTCTAC 1204
DB 1060 CAGAGCCCTACCCTGGGTCTGCGAGGCTGGCTGGGTGAGAGGAGTCTAC 1119
QY 1205 TGAGCTGCTTTGGTGGGACCAACCGGCGCACAGAAATGGCGGTGGGAGGAGTCTTCT 1264
DB 1120 TGAGCTGCTTTGGTGGGACCAACCGGCGCACAGAAATGGCGGTGGGAGGAGTCTTCT 1179
QY 1265 CAGGACCTCTCGCAAGACCGCTCTGGGAGAGAAATAAGCACTGGGAGATTGGAAGCACT 1324
DB 1180 CAGGACCTCTCGCAAGACCGCTCTGGGAGAGAAATAAGCACTGGGAGATTGGAAGCACT 1239
QY 1325 GCTAACATTTTGAATTTTCTTCTTAAATTTTAAAGATGTTAGTGTCTTAAAGCT 1384
DB 1240 GCTAACATTTTGAATTTTCTTCTTAAATTTTAAAGATGTTAGTGTCTTAAAGCT 1299
QY 1385 TTTATTTTTCGAACTTTTGAAGTCACTTCAATGAAGGTATTAATTTTACATAATAA 1444
DB 1300 TTTATTTTTCGAACTTTTGAAGTCACTTCAATGAAGGTATTAATTTTACATAATAA 1359
QY 1445 AAATGCACTCAATTT 1458
DB 1360 AAATGCACTCAATTT 1373
RESULT 2
BO067335
LOCUS 1134 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6767425 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5751735
ACCESSION BO067335
VERSION BO067335.1 GI:19896381
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1134)
AUTHORS NIH-MGC <http://mgi.nhl.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12784 row: m column: 16
High quality sequence stop: 607.
FEATURES
location/Qualifiers
1..1134
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5751735"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous Pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 57.7%; Score 841.6; DB 5; Length 1134;

	Best Local Similarity	93.8%; Pred. No. 2.4e-208;	Matches	966; Conservative	0; Mismatches	49; Indels	15; Gaps	8;
QY	6	GGAGATGGGATGTCCACAGATGATAGGCTCCTGGGATTTCAGAACCCAAAGACCAGCAGGAC	65					
Db	27	GGAGATGGGATGTCCACAGATGATAGGCTCCTGGGATTTCAGAACCCAAAGACCAGCAGGAC	86					
QY	66	TCCAGTCACTCTACCCCGAGCTCTCAGGACACAGCGCTCCCAAATCTTGAGTGACGTCCTC	135					
Db	87	TCCAGTCACTCTACCCCGAGCTCTCAGGACACAGCGCTCCCAAATCTTGAGTGACGTCCTC	146					
QY	126	ACCTCTGGTCTTTGCAGCACAAACAAGTGGGAATCACACCTCCAGACTCCCAAGCTT	195					
Db	147	ACCTCTGGTCTTTGCAGCACAAACAAGTGGGAATCACACCTCCAGACTCCCAAGCTT	206					
QY	186	CCACCCAGACTGGGCGCGGCCCTGCTCCATTTCAGCTGTGACAACTTCAGAGCCGCTG	245					
Db	207	CCACCCAGACTGGGCGCGGCCCTGCTCCATTTCAGCTGTGACAACTTCAGAGCCGCTG	266					
QY	246	TTGGCCCAAGCATGACAAAGACGTATGAATACTTCAGTAGTCTTGGAATAAAGTGAAAG	305					
Db	267	TTGGCCCAAGCATGACAAAGACGTATGAATACTTCAGTAGTCTTGGAATAAAGTGAAAG	326					
QY	306	TCCAGGGTTTTAAAATGGGCCACTTCTCTCCAGTCCCTCTGCAGCGTCTCGGCTCTG	365					
Db	327	TCCAGGGTTTTAAAATGGGCCACTTCTCTCCAGTCCCTCTGCAGCGTCTCGGCTCTG	386					
QY	366	GGCCCTGCCATCTCTGCTGCTCCCTGGGCTCGGCTCTGCTGCTGGTCAATCATCTGTG	425					
Db	387	GGCCCTGCCATCTCTGCTGCTCCCTGGGCTCGGCTCTGCTGCTGGTCAATCATCTGTG	446					
QY	426	TGTTGGATTCCAAAAATTCAAAATTCAGAGGACCTGGTGACCTCGAGAACAAGATTTTA	485					
Db	447	TGTTGGATTCCAAAAATTCAAAATTCAGAGGACCTGGTGACCTCGAGAACAAGATTTTA	506					
QY	486	GCAACTTCACCTCAAAACACT-GTGGCGAGATCCAGGCACATGACTTCCCAGGCGCAGC	544					
Db	507	GCAACTTCACCTCAAAACACTGGTGGCGAGATCCAGGCACATGACTTCCCAGGCGCAGC	566					
QY	545	TTGGAAGAAACGATAGCATCTCTGMAAGCTGA-GGTGGAGGGTTTCAAGCAGGAACGGCA	603					
Db	567	TTGGAAGAAACGATAGCATCTCTGMAAGCTGAGGGTGGAGGGTTTCAAGCAGGAACGGCA	626					
QY	604	GGCAGGGTATCTGAGCTCCAGGAACACATACGCAAGGCAACACTTAGGCCCATGTGCTC	663					
Db	627	GGCAGGGTATCTGAGCTCCAGGAACACACTAGCGAAGGCAACACTTAGGCCCATGTGCTC	686					
QY	664	CCACTGCCCATCTGTGTGTCTCCAGTTCAATCTGAAATGCTCCTGCGAGTCCAGCAGCT	723					
Db	687	CCACTGCCCATCTGTGTGTCTCCAGTTCAATCTGAAATGCTCCTGCGAGTTCAGCAGCT	746					
QY	724	GGTGAAGACCTGAAGAAACTGACCTGCCAGGTGGCTACTCTCAACAAATGCTCCAC	783					
Db	747	GGTGAAGACCTGAAGAAACTGACCTGCCAGGTGGCTACTCTCAACAAATGCTCCAC	806					
QY	784	TGAAGGACCTGTGCTGCCCGT-CAACTGGGTGGAGCACCAAGACAGCTGCTACTGTTCT	842					
Db	807	TGAAGGACCTGTGCTGCCCGT-CAACTGGGTGGAGCACCAAGACAGCTGCTACTGTTCT	866					
QY	843	CTCACTCTGGATGTCTT-GGGCGAGGCTGAGAAGTACTGCCAGCTTGAAGAGACGCCAC	901					
Db	867	CTCACTCTGGATGTCTTGGGGCCCAAGGCTGAGAAGTACTGCCAACTGAAGAGCGCCCA	926					
QY	902	CTGG--TGGTCACTAACT-CCAGGGAGGACAGAAATTTGTCAGAAATATCTAGG---	954					
Db	927	CCGGTGGGTCACTAACTCCAGGGAGGACCAAGAAATTTTGTCCGAAATAACTGGGGCT	986					
QY	955	-CTCCGATACACTCGATGGGCTCAGTGACCTCGAAGGAGCCTGG---AAGTGGGTGG	1010					
Db	987	CCGCAATACACCTGGGAAGGGCTTAAGTGACCTTGAAGGACCTTGAAAAATGGGGGG	1046					
QY	1011	ATGGAACAGA	1020					

```
QY 806 AACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCTCACTCTGGGATGCTCTGGGCC 865
Db |||||
QY 671 AACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCTCACTCTGGGATGCTCTGGGCC 612
Db |||||
QY 866 GAGGCTGAGAGTACTGCCAGCTGAAGAGGCCACCTGGTGTCTATCACTCCAGGGAG 925
Db |||||
QY 611 GAGGCTGAGAGTACTGCCAGCTGAAGAGGCCACCTGGTGTCTATCACTCCAGGGAG 552
QY 926 GAGCAAAATTTTCTCCAGAAATATCTAGGCTCGCATACACCTGGATGGCCCTCAGTGAC 985
Db |||||
QY 551 GAGCAAAATTTTCTCCAGAAATATCTAGGCTCGCATACACCTGGATGGCCCTCAGTGAC 492
QY 986 CTTGAAGAGCCCTGGAGTGGGTGGATGGAACAGACTATGCGACCGGCTTCCAGAACTGG 1045
Db |||||
QY 491 CTTGAAGAGCCCTGGAGTGGGTGGATGGAACAGACTATGCGACCGGCTTCCAGAACTGG 432
QY 1046 AAGCCAGGCCAGCCAGACACTGGCAGGGGACGGCTGGGTGGAGGCGAGACTGTGCT 1105
Db |||||
QY 431 AAGCCAGGCCAGCCAGACACTGGCAGGGGACGGCTGGGTGGAGGCGAGACTGTGCT 372
QY 1106 CACTTCCATCCAGACGGCAGGTGGAATGACAGCTCTGCCAGAGGCCCTTACCACCTGGGTC 1165
Db |||||
QY 371 CACTTCCATCCAGACGGCAGGTGGAATGACAGCTCTGCCAGAGGCCCTTACCACCTGGGTC 312
QY 1166 TCGAGGCTGGCTGGGTGAGCAGCCAGCAGGAGTCACTGAGCTGCCCTTTGGTGGGACC 1225
Db |||||
QY 1226 ACCGGCCACAGAAATGGCGGTGGAGAGGACTCTTCTCAGACCTCTCGCAGAGACGG 1285
Db |||||
QY 251 ACCGGCCACAGAAATGGCGGTGGAGAGGACTCTTCTCAGACCTCTCGCAGAGACGG 192
QY 1286 CTCCTGGAGAGAAATAGCAGCTGGGAGATTGGAAGCAGCTGCTAAACATTTTGAATTTTT 1345
Db |||||
QY 191 CTCCTGGAGAGAAATAGCAGCTGGGAGATTGGAAGCAGCTGCTAAACATTTTGAATTTTT 132
QY 1346 CTCCTTAAATTTTAAAGATGGTATAGTGTCTTAAAGCTTTTATTTTTTTCACACTTTT 1405
Db |||||
QY 1406 GAAAGTCACTTCATGAAGTATAATTTTACATAATAAAATGCACTCATTT 1458
Db |||||
QY 71 GAAAGTCACTTCATGAAGTATAATTTTACATAATAAAATGCACTCATTT 19

RESULT 4
AL546075
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

```
http://www.genoscope.cns.fr/cdna?s=CS0DI024G05QP1&c=8440.r.
Location/Qualifiers
1..827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI024YN10"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 50.5%; Score 736; DB 1; Length 827;
Best Local Similarity 91.1%; Pred. No. 8.6e-181;
Matches 827; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 5 AGGAGATGGGATGTCCAGATGATAGGGCTCCTGGGATTTTCAGACCCCAAGACCAGCAGCA 64
Db |||||
QY 1 AGGAGATGGGATGTCCAGATGATAGGGCTCCTGGGATTTTCAGACCCCAAGACCAGCAGCA 60
Db |||||
QY 65 CTCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGACGTC 124
Db |||||
QY 61 CTCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGACGTC 120
QY 125 CACCTCTGGTCTTTCAGACCAACCAAGTGGGAATCACACCTCCAGACCTCCACAGC 184
Db |||||
QY 121 CACCTCTGGTCTTTCAGACCAACCAAGTGGGAATCACACCTCCAGACCTCCACAGC 180
QY 185 TCCACCCAGACTGGGGCGGCTCTGCCTCAATTTTCAGCTGTGACAACTTCAGAGCCGT 244
Db |||||
QY 181 TCCACCCAGACTGGGGCGGCTCTGCCTCAATTTTCAGCTGTGACAACTTCAGAGCCGT 240
QY 245 GTTGCCCAAGCATGACAAGGACGATATGAAAATCTTCCAGTACTTTGGAGAATAAGGTGAA 304
Db |||||
QY 241 GTTGCCCAAGCATGACAAGGACGATATGAAAATCTTCCAGTACTTTGGAGAATAAGGTGAA 300
QY 305 GTCCAGGGGTTTAAATATGGGCGCACTCTCTCCAGTCCCTCCAGCGCTCCGCTCT 364
Db |||||
QY 301 GTCCAGGGGTTTAAATATGGGCGCACTCTCTCCAGTCCCTCCAGCGCTCCGCTCT 360
QY 365 GGGCCCTGCCATCTCTCTGCTGCTGGGCGCTCGGCGCTGCTGCTGCTGCTCATCTGT 424
Db |||||
QY 361 GGGCCCTGCCATCTCTCTGCTGCTGGGCGCTCGGCGCTGCTGCTGCTGCTCATCTGT 420
QY 425 GTGTTGGATTTCCAAAATTTCCAAAATTTCCAGAGGACCTGGTGACCTTGAGAACAGATTTT 484
Db |||||
QY 421 GTGTTGGATTTCCAAAATTTCCAAAATTTCCAGAGGACCTGGTGACCTTGAGAACAGATTTT 480
QY 485 AGCACTTTCACCTCAACACACTGTGGCGGAGATCCAGGCACTTGACTTCCAGGGCAGCAG 544
Db |||||
QY 481 AGCACTTTCACCTCAACACACTGTGGCGGAGATCCAGGCACTTGACTTCCAGGGCAGCAG 540
QY 545 TTGGAAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAACCGCAG 604
Db |||||
QY 541 TTGGAAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAACCGCAG 600
QY 605 GCAGGGGTATCTGAGCTTCCAGGAACACACTACGCAAGAGGCAACCTTAGGCCACTGTCCC 664
Db |||||
QY 601 G----- 601
QY 665 CACTGCCCATCTGTGTGTGTCCTCCAGTTCAATTTCTGAATGCTCTCGGAGTCCAGCAGCTG 724
Db |||||
QY 602 ----- 639
QY 725 GTGCAAGACCTGAAGAAATGACCTGCGCAGGTGGCTACTCTCAACAAACAATGCCCTCCACT 784
Db |||||
QY 640 GTGCAAGACCTGAAGAAATGACCTGCGCAGGTGGCTACTCTCAACAAACAATGCCCTCCACT 699
QY 785 GAAAGGACCTGTCGCGCCGCTCAACTGGGTGGAGCAACCAAGACAGCTGTCTACTGGTTCTCT 844
```

```

Db      700 GAAGGACCTGCTGCCCGTCAACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCT 759
Qy      845 CACTCTGGGATGCTCTGGGCGGAGGCTGAGAAGTACTGCGCAGTGAAGAACGCCCACTG 904
Db      760 CACTCTGGGATGCTCTGGGCGGAGGCTGAGAAGTACTGCGCAGTGAAGAACGCCCACTG 819
Qy      905 GTGGTCAAT 912
Db      820 GTGGTCAAT 827

RESULT 5
BI820020
LOCUS   BI820020      842 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 60303534F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176487 5',
mRNA sequence.
ACCESSION BI820020
VERSION   BI820020.1 GI:15931570
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 842)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          CDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLNL1439 Row: 1 Column: 24
          High quality sequence stop: 819.

FEATURES
source
1..842
Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5176487"
    /lab_host="DH10B"
    /clone_lib="NIH MGC 115"
    /notes="Organ: pooled brain, lung, testis; Vector:
    pCMV-SPORT6; Site: 1. NotI; Site 2: EcoRV (destroyed); RNA
    source anonymous pool of 6 male brains, age range 23-27; 1
    male lung, age 27; and 1 male testis, age 69. Library is
    oligo-dT primed and directionally cloned (EcoRV site is
    destroyed upon cloning). Average insert size 1.8 kb,
    insert size range 1-3 kb. Library is normalized and
    enriched for full-length clones and was constructed by C.
    Gruber (Invitrogen). Research Genetics tracking code
    021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 48.7%; Score 709.6; DB 4; Length 842;
Best Local Similarity 90.5%; Pred. No. 6.9e-174;
Matches 836; Conservative 0; Mismatches 4; Indels 84; Gaps 4;

Qy      1 GTTGAGGAGATGGGATGTCCAGATGATAGGGCTCTCTGGGATTTCCAGACCACAGCAGC 60
Db      1 GTTGAGGAGATGGGATGTCCAGATGATAGGGCTCTCTGGGATTTCCAGACCACAGCAGC 60
Qy      61 AGACTCTCAGTCACTTACCAGCTCTCCAGGACACAGCGGTCCCAACTCTGAGTGAC 120
Db      1 GTTGAGGAGATGGGATGTCCAGATGATAGGGCTCTCTGGGATTTCCAGACCACAGCAGC 60
Qy      61 AGACTCTCAGTCACTTACCAGCTCTCCAGGACACAGCGGTCCCAACTCTGAGTGAC 120
Db      1 GTTGAGGAGATGGGATGTCCAGATGATAGGGCTCTCTGGGATTTCCAGACCACAGCAGC 60
Qy      121 GTCCCACTCTGCTCTTGGAGCAGACACCAACGTTGGGAATCACACCTCCAGACCTCCCA 180
Db      121 GTCCCACTCTGCTCTTGGAGCAGACACCAACGTTGGGAATCACACCTCCAGACCTCCCA 180

```

```

Qy      181 CAGCTCAACCCAGAGACTGGGCGCGGCGCTGCCTCATTTTCAGCTGTGACAACTCTAGAG 240
Db      181 CAGCTCAACCCAGAGACTGGGCGCGGCGCTGCCTCATTTTCAGCTGTGACAACTCTAGAG 240
Qy      241 CCCTGTTGGCCCAAGCATGACAGGACGCTATGAAAATTTCCAGTACTTGGAGAAATAGGT 300
Db      241 CCCTGTTGGCCCAAGCATGACAGGACGCTATGAAAATTTCCAGTACTTGGAGAAATAGGT 300
Qy      301 GAAAGTCCAGGGGTTTTAAAAATGGGCCACTTCTCTCCAGTCCCTCTCTGAGGCTCTCCG 360
Db      301 GAAAGTCCAGGGGTTTTAAAAATGGGCCACTTCTCTCCAGTCCCTCTCTGAGGCTCTCTG 360
Qy      361 CTCTGGGCGCTGTCATCTCTGCTGCTCTGGGCGCTGCGCTGCTGCTGCTGCTGCTCATCAT 420
Db      361 CTCTGGGCGCTGTCATCTCTGCTGCTCTGGGCGCTGCGCTGCTGCTGCTGCTGCTCATCAT 420
Qy      421 CTGCTGTTGGATTCCAAAATTTCCAAAATTTCCAGAGGGACCTGCTGACCCCTGAGAACAGA 480
Db      421 CTGCTGTTGGATTCCAAAATTTCCAAAATTTCCAGAGGGACCTGCTGACCCCTGAGAACAGA 480
Qy      481 TTTTAGCAACTTCACTCAAACTCTGGCGGAGATCCAGGCACTGACTTTCCAGGGCAG 540
Db      481 TTTTAGCAACTTCACTCAAACTCTGGCGGAGATCCAGGCACTGACTTTCCAGGGCAG 540
Qy      541 CAGCTTGGAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAAACG 600
Db      541 CAGCTTGGAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAAACG 600
Qy      601 GCAGGACGAGGGTATCTGAGCTCCAGGACACACTACGCGAAGGACACACCTAGGCCACTG 660
Db      601 GCAGG-----
Qy      661 TCCCCACTGCCCCATCTGTGTGTGCCAGTTCATTTCTGAAATGCTCTCTGCGAGTCCAGCA 720
Db      606 -----CAGTTCAATCTGAAATGCTCTCTGCGAGTCCAGCA 639
Qy      721 GCTGTGCGAAGACCTGAGAAACTGACCTGCGCAGGTGGCTACTCTCTCAAACTGCTCTC 780
Db      640 GCTGTGCGAAGACCTGAGAAACTGACCTGCGCAGGTGGCTACTCTCTCAAACTGCTCTC 699
Qy      781 CACTGAAGGACCTGCTGCCCCGCTCAACTGGGTGGAGACCAAGACAGCTGCTACTGTT 840
Db      700 CACTGAAGGACCTGCTGCTG-CCCCGCTCAACTGGGTGGAGACCAAGACAGCTGCTACTGTT 758
Qy      841 CTCTCACTCT-GGGATGTCTCTGGGCGGAGGCTGAGAACTGCTGCCAGCTGAAGAACGCC 899
Db      759 CTCTCACTCTGGGATGTCTCTGGGCGGAGGCTGAGAACTGCTGCCAGCTGAAGAACGCC 818
Qy      900 ACCT-GGTGCTCATCAACTCCAGG 922
Db      819 ACCTGGGTGCTCATCAACTCCAGG 842

RESULT 6
BI821633
LOCUS   BI821633
DEFINITION 60303669F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5177584 5',
mRNA sequence.
ACCESSION BI821633
VERSION   BI821633.1 GI:15933183
KEYWORDS EST.
ORGANISM Homo sapiens (human)
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 829)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: Life Technologies, Inc.

```

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11442 row: j column: 17
 High quality sequence stop: 823.

FEATURES

```

Source
1. 829
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5177584"
/lab_host="DH10B"
/clone_lib="NIH MGC 115"
/notes="Organ: pooled brain, lung, testis; vector:
pCMV-SPOK16; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 48.4%; Score 706.2; DB 4; Length 829;
Best Local Similarity 90.5%; Pred. No. 5.3e-173;
Matches 821; Conservative 0; Mismatches 3; Indels 83; Gaps 3;

QY 10 ATGGGATGTCCAGATGATAGGCTCTGGGATTTTCAGACCCCAAGACAGGAGCTCCA 69
DB 1 ATGGGATGTCCAGATGATAGGCTCTGGGATTTTCAGACCCCAAGACAGGAGCTCCA 60

QY 70 GTCACCTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGACGTCCACCT 129
DB 61 GTCACCTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGACGTCCACCT 120

QY 130 CTGGTCTTTCAGACACCAAGCTGGGAATCACACCTTCAGACCTCCACAGCTCCAC 189
DB 121 CTGGTCTTTCAGACACCAAGCTGGGAATCACACCTTCAGACCTCCACAGCTCCAC 180

QY 190 CCAGACTGGGCGCGGCTGCTCCATTTTCAGCTGTGCACACTTCAGACCGCTGTGG 249
DB 181 CCAGACTGGGCGCGGCTGCTCCATTTTCAGCTGTGCACACTTCAGACCGCTGTGG 240

QY 250 CCCAAGCATGACAAGGACGTATGAAACTTCCAGTACTTTGGAGAATAAGGTGAAAGTCCA 309
DB 241 CCCAAGCATGACAAGGACGTATGAAACTTCCAGTACTTTGGAGAATAAGGTGAAAGTCCA 300

QY 310 GGGGTTTAAATGGGCGCACTTCTCTCAGTCCCTCTCGAGCGTCTCGCTCTGGGCC 369
DB 301 GGGGTTTAAATGGGCGCACTTCTCTCAGTCCCTCTCGAGCGTCTCTCTCTGGGCC 360

QY 370 CTGCCATCTCTCTCTCGGCGCTCGGCTGCTGCTGCTGCTGCTCATCTGCTGTGTGT 429
DB 361 CTGCCATCTCTCTCTCGGCGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGT 420

QY 430 TGGATTCCAAAATTCAAAATTTTCAGAGGACCTGTGACCCCTTGAGAACAGATTTTAGCAA 489
DB 421 TGGATTCCAAAATTCAAAATTTTCAGAGGACCTGTGACCCCTTGAGAACAGATTTTAGCAA 480

QY 490 CTTCACTTCAACACTGTGGCGAGATCCAGGCACTGACTTCCAGGCGAGCAGCTTGA 549
DB 481 CTTCACTTCAACACTGTGGCGAGATCCAGGCACTGACTTCCAGGCGAGCAGCTTGA 540

QY 550 AGAAACGATAGCATCTCTGAAAGCTCAGGTGGAGGGTTTCAAGCAGGAACCGCAGCAGG 609
DB 541 AGAAACGATAGCATCTCTGAAAGCTCAGGTGGAGGGTTTCAAGCAGGAACCGCAGCAGG 609

QY 610 GGTATCTGAGTCCAGGAACACACTACGAGGAAGGCACACTTAGGCCACTGTCCCACTG 669

```

```

Db 597 ----- 596
QY 670 CCCATCTGTGTGTCGCCAGTTTCATTCTGAAATGCTCTCGAGTCCAGCAGCTGTGCA 729
Db 597 -----CAGTTCATTCTGAAATGCTCTCTCGAGTCCAGCAGCTGTGCA 639
QY 730 AGACTGTAAGAACTGACCTGCCAGGTGGTCTCTCAACAACTGCTCCACTGAAGG 789
Db 640 AGACTGTAAGAACTGACCTGCCAGGTGGTCTCTCAACAACTGCTCCACTGAAGG 699
QY 790 GACCTGTGCCCCGTCAACTGGGTGGAGCACCAAGAGCTGCTACTGTGTTCTCTCACTC 849
Db 700 GACCTGTGCCCCGTCAACTGGGTGGAGCACCAAGAGCTGCTACTGTGTTCTCTCACTC 759
QY 850 TGGGATGTCTGGGCGCGAGG-CTGAGAAGTACTGCCAGCTGAAGAACGCCACCTGTGG 908
Db 760 TGGGATGTCTGGGCGCGAGGCTGAGAAGTACTGCCAGCTGAAGAACG-CCACCTGTGG 818
QY 909 TCATCAA 915
Db 819 TCATCAA 825

RESULT 7
LOCUS BU621880 678 bp mRNA linear EST 23-SEP-2002
DEFINITION UI-H-DF0-bev-j-20-0-UI.s1 NCI CGAP_DF0 Homo sapiens cDNA clone
ACCESSION UI-H-DF0-bev-j-20-0-UI 3', mRNA sequence.
VERSION BU621880
KEYWORDS BU621880.1 GI:23288095
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 678)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-22, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

```

FEATURES

```

Location/Qualifiers
1. 678
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-bev-j-20-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DF0"
/notes="Organ: Bone; Vector: p77T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DF0 is a cDNA library containing the following
tissues: Subchondral Bone. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into p77T3-Pac
vector. The oligonucleotide used to prime the synthesis of

```

first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTTAAAGCGTC.
TAG TISSUE=chondrosarcoma bone
TAG LIB=UI-H-DF0
TAG_SEQ=GTAAAGCGTC

ORIGIN

Query Match 45.2%; Score 658.4; DB 5; Length 678;
Best Local Similarity 99.8%; Pred. No. 1.6e-160;
Matches 659; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 799 CCCGCTCAACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCTCACTCTGGATGTC 858
DB 678 CCTGTCACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCTCACTCTGGATGTC 619
QY 859 CTGGCCGAGGCTGAGAAGTACTGCCAGCTGAAGACGCCACCTCGTGGTGCATCAACTC 918
DB 618 CTGGCCGAGGCTGAGAAGTACTGCCAGCTGAAGACGCCACCTCGTGGTGCATCAACTC 559
QY 919 CAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCCT 978
DB 558 CAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCCT 499
QY 979 CAGTGACCTGAAGGAGCTGGAGTGGGTGGATGGACAGACTATCGACCGCTTCCA 1038
DB 498 CAGTGACCTGAAGGAGCTGGAGTGGGTGGATGGACAGACTATCGACCGCTTCCA 439
QY 1039 GAACTGGAAGCCAGGCGCAGACAGCTGCGAGGGGACCGGCTGGGTGGAGCGAGGA 1098
DB 438 GAACTGGAAGCCAGGCGCAGACAGCTGCGAGGGGACCGGCTGGGTGGAGCGAGGA 379
QY 1099 CTGTGCTCACTTCCATCCAGACGGCAGGTGGAATGACGAGCTCTGCCAGAGGCCCTACCA 1158
DB 378 CTGTGCTCACTTCCATCCAGACGGCAGGTGGAATGACGAGCTCTGCCAGAGGCCCTACCA 319
QY 1159 CTGGGTCTGAGGCTGGCTGGGTGACAGCAGGAGGAGTCACTGAGCTGGCTTTGG 1218
DB 318 CTGGGTCTGAGGCTGGCTGGGTGACAGCAGGAGGAGTCACTGAGCTGGCTTTGG 259
QY 1219 TGGGACCAACCGGCGCAGAGAAATGGCGGTGGAGGAGGACTCTTCTCAGCACTCTCTCGC 1278
DB 258 TGGGACCAACCGGCGCAGAGAAATGGCGGTGGAGGAGGACTCTTCTCAGCACTCTCTCGC 199
QY 1279 AAGACCGCTCTGGGAGAGAAATAGCACTGGAGATTGGAAGCACTCTTAACATTTTGA 1338
DB 198 AAGACCGCTCTGGGAGAGAAATAGCACTGGAGATTGGAAGCACTCTTAACATTTTGA 139
QY 1339 TTTTCTCTTTAAATTTTAAAGATGGTATAGTGTCTTAAGCTTTTATTTTTC 1398
DB 138 TTTTCTCTTTAAATTTTAAAGATGGTATAGTGTCTTAAGCTTTTATTTTTC 79
QY 1399 AACTTTTGAAGTCACTCTCATGAGGTATATTTTACATAATAAATAAGCACTATT 1458
DB 78 AACTTTTGAAGTCACTCTCATGAGGTATATTTTACATAATAAATAAGCACTATT 19

RESULT 8

CA412958/c
LOCUS CA412958 674 bp mRNA linear EST 07-NOV-2002
DEFINITION UI-H-E20-bao-b-08-0-UI.s1 NCI_CGAP_Ch1 Homo sapiens cDNA clone
UI-H-E20-bao-b-08-0-UI 3', mRNA sequence.
ACCESSION CA412958
VERSION CA412958.1 GI:24775609
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 674)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-femail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of
Orthopaedics
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-22, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA-Yes.

FEATURES

source

1. .674

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-E20-bao-b-08-0-UI"
/tissue_type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Ch1"
/note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI CGAP Ch1 is a cDNA library containing the following
tissue(s): Chondrosarcoma Grade II. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TGATCAGCT.
TAG TISSUE=grade-2-chondrosarcoma
TAG LIB=UI-H-E20
TAG_SEQ=ATCTAATATG

ORIGIN

Query Match 45.0%; Score 656; DB 6; Length 674;
Best Local Similarity 100.0%; Pred. No. 6.6e-160;
Matches 656; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 803 GTCAACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTCCTGG 862
DB 674 GTCAACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTCCTGG 615
QY 863 GCCGAGGCTGAGAAGTACTGCCAGCTGAAGAACGCCACCTGGTGGTCATCAACTCCAGG 922
DB 614 GCCGAGGCTGAGAAGTACTGCCAGCTGAAGAACGCCACCTGGTGGTCATCAACTCCAGG 555
QY 923 GAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGT 982
DB 554 GAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGT 495
QY 983 GACCTCTGAAGGAGCTGGAGTGGGTGGATGGACAGACTATGCGACCGGCTTCCAGAAC 1042
DB 494 GACCTCTGAAGGAGCTGGAGTGGGTGGATGGACAGACTATGCGACCGGCTTCCAGAAC 435
QY 1043 TGGAGCCAGGCCAGCCAGACGACTGGCAGGGGACCGGGCTGGGTGGAGGCGAGGACTGT 1102
DB 434 TGGAGCCAGGCCAGCCAGACGACTGGCAGGGGACCGGGCTGGGTGGAGGCGAGGACTGT 375
QY 1103 GCTCACTTCCATCCAGACGCGAGTGAATGACGACTCTGCCAGAGGCCCTACCACTGG 1162
DB 374 GCTCACTTCCATCCAGACGCGAGTGAATGACGACTCTGCCAGAGGCCCTACCACTGG 315
QY 1163 GTCTGGGAGGCTGGCTGGGTGACACCGCAGGAGAGTCACTGAGCTGCCCTTTGGTGGG 1222

Db	314	GTCTCGAGGCTGGCCCTGGGTGAGACCAAGGAGAGTCACTGAGTGCCTTTGGTGGG	255
Qy	1223	ACCACCCGCCACAGAAATGGCGTGGGAGGAGGACTCTTCTCAGACCTCTCGCAAGA	1282
Db	254	ACCACCCGCCACAGAAATGGCGTGGGAGGAGGACTCTTCTCAGACCTCTCGCAAGA	195
Qy	1283	CGCTCTGGGAGAGAAATAGCACTGGGAGATTGGAAGCACCTGCTAAACATTTTGAATTT	1342
Db	194	CGCTCTGGGAGAGAAATAGCACTGGGAGATTGGAAGCACCTGCTAAACATTTTGAATTT	135
Qy	1343	TTTCTCTTAATTTTAAAGATGGTATAGTGTCTTAAGCTTTTATTTTTCCACT	1402
Db	134	TTTCTCTTAATTTTAAAGATGGTATAGTGTCTTAAGCTTTTATTTTTCCACT	75
Qy	1403	TTTGAAGTCAACTTCATGAGGTATAATTTTACATAATAAAATGCACTCATTT	1458
Db	74	TTTGAAGTCAACTTCATGAGGTATAATTTTACATAATAAAATGCACTCATTT	19
RESULT 9			
BI838214			
LOCUS			
DEFINITION			
603083132F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:522622 5',			
mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-remail.nih.gov			
Tissue Procurement: Life Technologies, Inc.			
cDNA Library Preparation: Life Technologies, Inc.			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: L1AM11559 row: 0 column: 07			
High quality sequence stop: 719.			
FEATURES			
source			
1. .721			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:522622"			
/lab_host="DH10B"			
/clone_lib="NIH MGC 120"			
/note="Organ: pooled pancreas and spleen; Vector:			
pCMV-SPT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA			
source anonymous Pool of spleen and pancreas from 28 yo			
male. Library is oligo-dT primed and directionally cloned			
(EcoRV site is destroyed upon cloning). Average insert			
size 1.5 kb, insert size range 1-2.5 kb. Library is			
normalized and enriched for full-length clones and was			
constructed by C. Gruber (Invitrogen). Research Genetics			
tracking code 025. Note: this is a NIH_MGC Library."			
ORIGIN			
Query Match 45.0%; Score 655.8; DB 4; Length 721;			
Best Local Similarity 99.6%; Pred. No. 7.6e-160;			
Matches 668; Conservative 0; Mismatches 2; Indels 1; Gaps 1;			
Qy	1	GTTCAGGAGTGGATGTCCTCCAGATGATAGGGCTCTCGGATTTACAGCCAGACCAGC	60
Db	41	GTTCAGGAGTGGATGTCCTCCAGATGATAGGGCTCTCGGATTTACAGCCAGACCAGC	100

Qy	61	AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTCA	120
Db	101	AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTCA	160
Qy	121	GTCCACCTCTGGTCTTGGAGCACCAACAGTGGGAATCACACCTCCAGACCTCCCA	180
Db	161	GTCCACCTCTGGTCTTGGAGCACCAACAGTGGGAATCACACCTCCAGACCTCCCA	220
Qy	181	CAGCTCCACCCAGACTGGGCGCGCGCTCTCCATTTTCAGCTGTGCAACACCTCAG	240
Db	221	CAGCTCCACCCAGACTGGGCGCGCGCTCTCCATTTTCAGCTGTGCAACACCTCAG	280
Qy	241	CCGTGTTGGCCCAAGCATGACAAAGACGTATGAAACATTCAGTACTTTGGGAATAA	300
Db	281	CCGTGTTGGCCCAAGCATGACAAAGACGTATGAAACATTCAGTACTTTGGGAATAA	340
Qy	301	GAAAGTCCAGGGGTTTAAATAATGGGCACTTCTCTCCAGTCCCTCCGACGCTCTCG	360
Db	341	GAAAGTCCA - GGGTTTAAATAATGGGCACTTCTCTCCAGTCCCTCCGACGCTCTCG	399
Qy	361	CTCTGGGCGCTGCCATCTCTCTGCTGTCTGGGCTCGGCTCTGCTGCTGGTGCATCAT	420
Db	400	CTCTGGGCGCTGCCATCTCTCTGCTGTCTGGGCTCGGCTCTGCTGCTGGTGCATCAT	459
Qy	421	CTGTGTGGTGGATTCCTCAAAATTTCCAAATTTCCAGAGGACCTGTGACCTCGAGAAC	480
Db	460	CTGTGTGGTGGATTCCTCAAAATTTCCAAATTTCCAGAGGACCTGTGACCTCGAGAAC	519
Qy	481	TTTTRAGCAACTTCACCTCAAACTGTGGCGGAGATCCAGGCACTGCTCCAGGGCAG	540
Db	520	TTTTRAGCAACTTCACCTCAAACTGTGGCGGAGATCCAGGCACTGCTCCAGGGCAG	579
Qy	541	CAGCTTTGAAGAAACGATAGCATCTCTCAAACTGTGGTGGAGGGTTTCAAGCAGGAACG	600
Db	580	CAGCTTTGAAGAAACGATAGCATCTCTCAAACTGTGGTGGAGGGTTTCAAGCAGGAACG	639
Qy	601	GCAGGACGGGTATCTGAGTCTCCAGGAACACACTACGAGGAAGGCACACCTAGGCCACTG	660
Db	640	GCAGGACGGGTATCTGAGTCTCCAGGAACACACTACGAGGAAGGCACACCTAGGCCACTG	699
Qy	661	TCCCCCACTGCC 671	
Db	700	TCCCCCACTGCC 710	
RESULT 10			
EUG87211/c			
LOCUS			
DEFINITION			
UI-CF-EC1-adv-g-09-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone			
UI-CF-EC1-adv-g-09-0-UI 3', mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
Contact: McCray, PB			
University of Iowa			
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA			
Tel: 319 356 4866			
Fax: 319 356 7171			
Email: paul-mccray@uiowa.edu			
Tissue Procurement: Dr. M. J. Welsh, University of Iowa			
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa			

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

Location/Qualifiers
 1. 710
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-EC1-adv-g-09-0-UI"
 /tissue_type="Lung"
 /dev_stage="Adult and Fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-EC1"
 /notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: Ecor I; Site 2: Not I;
 UI-CF-EC1 is a normalized cDNA library containing the
 following tissue(s): Normal lung from adult and from fetal
 day 64, day 87, week 19 and week 42. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an Ecor I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 AAGTGGCTTAC.
 TAG TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
 and 380-383
 TAG LIB=UI-CF-EC1
 TAG_SEQ=AAGTGGCTTAC"

ORIGIN

Query Match 44.9%; Score 654.8; DB 5; Length 710;
 Best Local Similarity 99.3%; Pred. No. 1.4e-159;
 Matches 667; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
 787 AGGACCTGCTGCCCCGCTCAACTGGGTGGAGCACCAGACAGCTGCTACTGGTTCTCTCA 846
 709 AAGGACCTGCTGCCCCGCTCAACTGGGT-GAGCACCAGACAGCTGCTACTGGTTCTCTCA 651
 847 CTCTGGATGCTCTGGGCCGAGCTGAGAAGTACTGCCAGCTGAAGAACGCCACCTGGT 906
 650 CTCTGGATGCTCTGGGCCGAGCTGAGAAGTACTGCCAGCTGAAGAACGCCACCTGGT 591
 907 GGTCAATCACTCCAGGAGGAGCAGAAATTTGTCCAGAAATATCTAGGCTCCGCATACAC 966
 590 GGTCAATCACTCCAGGAGGAGCAGAAATTTGTCCAGAAATATCTAGGCTCCGCATACAC 531
 967 CTGGATGGGCTCAGTGACCTGAAGAGCTGGAGTGGTGGTGGTGAACAGACTATGC 1026
 530 CTGGATGGGCTCAGTGACCTGAAGAGCTGGAGTGGTGGTGGTGAACAGACTATGC 471
 1027 GACCGGCTTCAGAACTGGAAGCCAGCCAGCAGCTGCGAGGGGACCGGGCTGGG 1086
 470 GACCGGCTTCAGAACTGGAAGCCAGCCAGCAGCTGCGAGGGGACCGGGCTGGG 411
 1087 TGGAGGCGAGGACTGTGCTCACTTCCATCCAGAGCGGAGGTGGAATGACGACGCTCTGCCA 1146
 410 TGGAGGCGAGGACTGTGCTCACTTCCATCCAGAGCGGAGGTGGAATGACGACGCTCTGCCA 351
 1147 GAGGCCCTTACCTGGGCTCTGCAGGCTGGCGCTGGGTCAACAGCCAGGAGAGTCACTG 1206
 350 GAGGCCCTTACCTGGGCTCTGCAGGCTGGCGCTGGGTCAACAGCCAGGAGAGTCACTG 291
 1207 AGCTGCTTTGGTGGGACCAACCGGCCACAGAAATGCGGCTGGGAGGAGGACTCTTCTCA 1266
 290 AGCTGCTTTGGTGGGACCAACCGGCCACAGAAATGCGGCTGGGAGGAGGACTCTTCTCA 231

QY 1267 CGACCTCTCGCAAGACCGCTCTGGGAGAGAAATAGCACTGGAGATTCGAAGCACTGC 1326
 DB 230 CGACCTCTCGCAAGACCGCTCTGGGAGAGAAATAGCACTGGAGATTCGAAGCACTGC 171
 QY 1327 TAACATTTTGAATTTTCTCTTTTAAATTTTAAAAAGATGTTATAGTGTCTTTAAAGCTTT 1386
 DB 170 TAACATTTTGAATTTTCTCTTTTAAATTTTAAAAAGATGTTATAGTGTCTTTAAAGCTTT 111
 QY 1387 TATTTTTTTTCCAACTTTTGAAGTCAACTTCATGAAGGTATATATTTTACATAATAAAA 1446
 DB 110 TATTTTTTTTCCAACTTTTGAAGTCAACTTCATGAAGGTATATATTTTACATAATAAAA 51
 QY 1447 ATGCACCTCATTT 1458
 DB 50 ATGCACCTCATTT 39

RESULT 11

BI820664

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI820664 774 bp mRNA linear EST 04-OCT-2001
 603034446F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175683 5',
 mRNA sequence.
 BI820664
 BI820664.1 GI:15932214
 EST.
 Homo sapiens (human)
 Homo sapiens
 Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 774)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLAM1437 row: k column: 12
 High quality sequence stop: 769.

FEATURES

source

Location/Qualifiers
 1. 774
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5175683"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /notes="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPOK6; Site 1: NotI; Site 2: EcorV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcorV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 44.9%; Score 654; DB 4; Length 774;
 Best Local Similarity 96.3%; Pred. No. 2.3e-159;
 Matches 693; Conservative 0; Mismatches 15; Indels 10; Gaps 2;
 750 GCCAGGTGGTCTCTCAACAAAT-----GCTCCACTGAAGGACCTGCTGCC 800
 1 GCCAGGTGGTCTCTCAACAAATGGTGAGGAGCTCTCACTGAAGGACCTGCTGCC 60

QY	801	CCGTCAACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTCT	860
Db	61	CCGTCAACTGGGTGGAGCACCAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTCT	120
QY	861	GGGCCGAGCTGAGAAGTACTGCCAGCTGAAGAACGCCCACTGTGTGTCATCAACTCCA	920
Db	121	GGGCCGAGCTGAGAAGTACTGCCAGCTGAAGAACGCCCACTGTGTGTCATCAACTCCA	180
QY	921	GGGAGGAGCAGAAATTTGTCCAGAAATATCTAGGCTCCGCATACACCTTGGATGGGCTCA	980
Db	181	GGGAGGAGCAGAAATTTGTCCAGAAATATCTAGGCTCCGCATACACCTTGGATGGGCTCA	240
QY	981	GTGACCTCAAGGAGCTCGGAAGTGGGTGGATGGAAACAGACTATGCGACCGCTTCCAGA	1040
Db	241	GTGACCTCAAGGAGCTCGGAAGTGGGTGGATGGAAACAGACTATGCGACCGCTTCCAGA	300
QY	1041	ACTGGAAGCCAGCCAGCAGACGACTGCGAGGGGACCGGGCTGGGTGGAGCGGAGACT	1100
Db	301	ACTGGAAGCCAGCCAGCAGACGACTGCGAGGGGACCGGGCTGGGTGGAGCGGAGACT	360
QY	1101	GTGCTCACTTCCATCCAGACGGCAGGTGGAATGACAGCTCTGCCAGAGGCGCTTACCCT	1160
Db	361	GTGCTCACTTCCATCCAGACGGCAGGTGGAATGACAGCTCTGCCAGAGGCGCTTACCCT	420
QY	1161	GGGTCTGCGAGGCTGGCCCTGGGTGACAGCCAGCAGGAGTCACTGAGCTGCTTTGGTG	1220
Db	421	GGGTCTGCGAGGCTGGCCCTGGGTGACAGCCAGCAGGAGTCACTGAGCTGCTTTGGTG	480
QY	1221	GGACACCCGGCCACAGAAATGGCGTGGAGAGGACTTCTTCAAGACCTCTCTCGCAA	1280
Db	481	GGACACCCGGCCACAGAAATGGCGTGGAGAGGACTTCTTCAAGACCTCTCTCGCAA	540
QY	1281	GACCGCTCTGGGAGAGAAATAGCACTGGAGATTTGGAAGCACTGTACATTTTGAATT	1340
Db	541	GACCGCTCTGGGAGAGAAATAGCACTGGAGATTTGGAAGCACTGTACATTTTGAATT	600
QY	1341	TTTTCTCTTTAATTTTAAAGATGGTATAGTGTCTTTAAGCTTTATTTTTTCCAA	1400
Db	601	TTTTCTCTTTAATTTTAAAGATGGTATAGTGTCTTTAAGCTTTATTTTTTCCAA	660
QY	1401	CTTTTGA-AGTCAACTCATGAGGTATATTTTACATATATATAATGCACTCATTT	1457
Db	661	CTTTTGAAGTCACTTCAATGAGGTATATTTTACATATATAATGCACTCATTT	718
RESULT 12			
BI83888			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

FEATURES
source

Location/Qualifiers
1. 728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5226471"
/lab_host="DH10B"
/clone_lib="NIH_MGC 120"
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN

Query Match	44.5%;	Score 649;	DB 4;	Length 728;
Best Local Similarity	99.4%;	Pred. No. 4.5e-158;		
Matches 682;	Conservative 0;	Mismatches 1;	Indels 3;	Gaps 3;
QY	3	TGAGGAGATGGGATGTCCAGATGATAGGGCTCTCTGGGATTTCCAGACCCAAAGACAGCAG	62	
Db	45	TGAGGAGATGGGATGTCCAGATGATAGGGCTCTCTGGGATTTCCAGACCCAAAGACAGCAG	104	
QY	63	GACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAGCT	122	
Db	105	GACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAGCT	164	
QY	123	CCCACTCTGTGCTTT-GCAGCACAAACCAAGCTGGGAATCACACCTCCAGACCTCCAC	181	
Db	165	CCCACTCTGTGCTTTGGCAGCACAAACCAAGCTGGGAATCACACCTCCAGACCTCCAC	224	
QY	182	AGCTCCACCCAGACTGGGCGCGGCTGCTCTCAATTTACGCTGTGACAACTCCAGAGC	241	
Db	225	AGCTCCACCCAGACTGGGCGCGGCTGCTCTCAATTTACGCTGTGACAACTCCAGAGC	284	
QY	242	CGTGTGGCCCAAGCATGACAGGACGATGATAAACTTCCAGTACTTGGAGAAATAGGTG	301	
Db	285	CGTGTGGCCCAAGCATGACAGGACGATGATAAACTTCCAGTACTTGGAGAAATAGGTG	344	
QY	302	AAAGTCCAGGGGTTTAAATAATGGGCACTTCTCTCCAGTCCCTCTCCAGCGTCTCCGC	361	
Db	345	AAAGTCCA-GGGTTTAAATAATGGGCACTTCTCTCCAGTCCCTCTCCAGCGTCTCCGC	403	
QY	362	TCTGGGCGCTGCCATCTCTGCTGTCTGGGCTTGGGCTTGGCTGTGCTGTGTCATATC	421	
Db	404	TCTGGGCGCTGCCATCTCTGCTGTCTGGGCTTGGGCTTGGCTGTGCTGTGTCATATC	463	
QY	422	TGTGTGGTTGGATTTCCAAAATTTCCAAAATTTCCAGAGGGACCTGGTGACCTGAGACAGAT	481	
Db	464	TGTGTGGTTGGATTTCCAAAATTTCCAAAATTTCCAGAGGGACCTGGTGACCTGAGACAGAT	523	
QY	482	TTTAGCAACTTCACTCAAAACACTGTGGCGAGATCCAGGCACTGACTTCCAGGGCAGC	541	
Db	524	TTTAGCAACTTCACTCAAAACACTGTGGCGAGATCCAGGCACTGACTTCCAGGGCAGC	583	
QY	542	AGCTTGGAAAGAACGATAGCATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAAACGG	601	
Db	584	AGCTTGGAAAGAACGATAGCATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAAACGG	643	
QY	602	CAGGAGGGGTATCTGAGCTCCAGGACACACTAGGAGAGGGCACACTTAGGCCACTGT	661	
Db	644	CAGGAGGGGTATCTGAGCTCCAGGAAACACTAGGAGAGGGCACACTTAGGCCACTGT	703	
QY	662	CCCCACTGCCCATCTGTGTGTGTCCTC 687		
Db	704	CCCCACTG-CCATCTGTGTGTGTCCTC 728		

Db	483	AACTGGGTGGATGGAAACAGACTATGCGACGGGCTTCCAGAACTGGAGACGAGCCACGCA	424
Qy	1061	GACACCTGGCAGGGGCACGGGCTGGGTGGAGGGGAGGACTGTGCTCACTTCCATCCAGAC	1120
Db	423	GACGACTGGCAGGGGCACGGGCTGGGTGGAGGGGAGGACTGTGCTCACTTCCATCCAGAC	364
Qy	1121	GGCAGGTGGAAATGACGACGCTCTGCCAGAGGCCCTACCACTGGGTCTGCGAGGCTGGCGCT	1180
Db	363	GGCAGGTGGAAATGACGACGCTCTGCCAGAGGCCCTACCACTGGGTCTGCGAGGCTGGCGCT	304
Qy	1181	GGTCAGACCAGCAGGAGAGTCACTGAGCTGCTTTGGTGGGACCAACCCGGCCACAGAAA	1240
Db	303	GGTCAGACCAGCAGGAGAGTCACTGAGCTGCTTTGGTGGGACCAACCCGGCCACAGAAA	244
Qy	1241	TGCGGGTGGGAGGAGGACTCTTCTCACGACCTCTCGCAAGACCGCTCTGGGAGAGAAAT	1300
Db	243	TGCGGGTGGGAGGAGGACTCTTCTCACGACCTCTCGCAAGACCGCTCTGGGAGAGAAAT	184
Qy	1301	AAGCACTGGGAGATTTGGAAGCACTGCTAAACATTTTGAATTTTTTCTCTTAATTTTAAA	1360
Db	183	AAGCACTGGGAGATTTGGAAGCACTGCTAAACATTTTGAATTTTTTCTCTTAATTTTAAA	124
Qy	1361	AAGATGGTATAGTGTCTTTAAGCTTTTATTTTTTTTCCAACTTTTGAAGTCAACTTCAT	1420
Db	123	AAGATGGTATAGTGTCTTTAAGCTTTTATTTTTTTTCCAACTTTTGAAGTCAACTTCAT	64
Qy	1421	GAAGGTATATATTTTTTACATAATAAAATGCATCATTT	1458
Db	63	GAAGGTATATATTTTTTACATAATAAAATGCATCATTT	26
RESULT 14			
BI905917			
LOCUS	63083080F1 NIH_MGC_118	Homo sapiens	CDNA clone IMAGE:5212290 5',
DEFINITION		mRNA sequence.	
ACCESSION	BI905917		
VERSION	BI905917.1	GI:16168536	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	NIH-MGC	http://mgs.nci.nih.gov/	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) CDNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLAM1532 row: p column: 19 High quality sequence stop: 632.		
FEATURES			
source			
	1..634	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5212290"		
	/issue_type="leukocyte"		
	/lab_host="DH10B"		
	/clone_lib="NIH_MGC_118"		
	/notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range		


```
QY 599 CGGCGGCGGGGTATCTGAGCTCCAGGAACACACTAGCGAGGCGACACCTAGGCCAC 658
Db      |||||
624 CGGAGG----- 630
QY 659 TGTCCCCACTGCCCATCTGTGTGTGCCAGTTCAATTCTGAAATGCTCTCGGAGTCCAG 718
Db      |||||
631 -----CAGTTCAATCTGAAATGCTCTCGGAGTCCAG 662
QY 719 CAGCTGTGTC-AAGACCTGAAGAAACTGACCTGCCAGGTGGCTACTCTCAACAACAAT-- 775
Db      |||||
663 CAGCTGTGCAAGAACCTGAAGAAACTGACCTGCCAGGTGGCTACTCTCAACAACAATGG 722
QY 776 -----GCCTCCACTGAAGGGACCTGCTGCCCGTCAACTGGGTGGAGCACCAAGACAG 828
Db      |||||
723 TGAGGAAGCCTCCACTGAAGGGACCTGCTGCCCTGTCAACTGGGTGGAGCACCAAGACAG 782
QY 829 CTGCTACTGGTTC-TCTCACTCTGGGATGTC 858
Db      |||||
783 CTGCTACTGGTTCCTCTCACTCTGGGATGTC 813
```

Search completed: May 27, 2005, 23:52:45
Job time : 5121 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 19:12:00 ; Search time 6433 Seconds
(without alignments)
10982.078 Million cell updates/sec

Title: US-10-829-107-3
Perfect score: 1458
Sequence: 1 gttgaggaatggatgtcc.....taataaaatgcactcattt 1458

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1458	100.0	1458	6	BD135981 Isolated
2	1458	100.0	1458	6	AR561599 Sequence
3	1446.6	99.2	1769	9	BC039011 Homo sapi
4	1218	83.5	1344	9	D50532 Homo sapien
5	1108	76.0	1370	6	BD135983 Isolated
6	1108	76.0	1370	6	AR561601 Sequence
7	857.4	58.8	1858	9	BC027858 Homo sapi
8	606.6	41.6	661	6	AX067341 Sequence
9	438	30.0	1358	10	RAT03GSLC
10	420.8	28.9	172095	2	AC026834 Homo sapi
11	420.8	28.9	176813	2	AC107895 Homo sapi
12	420.8	28.9	198821	9	AC120057 Homo sapi
13	420.8	28.9	220581	2	AC015918 Homo sapi
14	404.6	27.8	1414	10	S36676 galactose a
15	403	27.6	1353	10	BC014811 Mus muscu
16	399.8	27.4	1520	10	AY103461 Mus muscu
17	382	26.2	1196	9	BC032130 Homo sapi
18	382	26.2	1277	6	E12702 cDNA encodi
19	382	26.2	1277	6	AR270623 Sequence

20	382	26.2	1277	6	AX409583	Sequence
21	382	26.2	1277	9	HUMASGPR1	
22	382	26.2	1657	9	AB070933	M10058 Human asial
23	374.6	25.7	873	9	CR542052	CR542052 Homo sapi
24	370	25.4	1223	6	CQ725393	CQ725393 Sequence
25	344.2	23.6	853	11	BV208440	BV208440 CLECF14
26	344.2	23.6	853	11	BC022106	BC022106 Mus muscu
27	335.6	23.0	1232	10	MUSMHL1	D13517 Mouse mRNA
28	334	22.9	991	10	MMU09362	U09362 Mus musculu
29	329	22.6	1309	10	MMU08372	E12703 cDNA encodi
30	321.4	22.0	1300	6	E12703	E12703 cDNA encodi
31	321.4	22.0	1300	6	AX411087	AX411087 Sequence
32	321.4	22.0	1300	9	HSLH2MR	X55283 Human L-H2
33	321.4	22.0	1344	9	BC017251	BC017251 Homo sapi
34	320.2	22.0	971	10	RATASGP2	M21770 Rat asialog
35	310.4	21.3	361	6	AX897915	AX897915 Sequence
36	310.4	21.3	361	6	BD033448	BD033448 Sequence
37	307.8	21.1	921	9	AF529374	AF529374 Homo sapi
38	290.6	19.9	879	9	H5U97197	U97197 Homo sapien
39	283.8	19.5	1308	6	CQ725392	CQ725392 Sequence
40	282.8	19.4	1309	6	AX409588	AX409588 Sequence
41	282.8	19.4	1309	6	AX774853	AX774853 Sequence
42	282.8	19.4	1309	9	HUMASGPR2	M11025 Human asial
43	273.4	18.8	1382	9	HSLH2BMR	X55284 Human H2' m
44	267	18.3	428	10	AY643476	AY643476 Marmota m
45	263	18.0	1470	10	AF230645	AF230645 Rattus no

ALIGNMENTS

RESULT 1

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD135981 Isolated dentritic cell membrane protein genes.
1458 bp DNA linear PAT 18-SEP-2002

BD135981 Isolated dentritic cell membrane protein genes.

BD135981.1 GI:23230926

JP 2002509438-A/2.

unidentified

unclassified.

1 (bases 1 to 1458)

Valladeau,J., Ravel,O., Bates,E.E.M., Ford,J., Saeland,S. and

Lebecque,S.J.E.

Isolated dentritic cell membrane protein genes

Patent: JP 2002509438-A 2 26-MAR-2002;

SCHERING CORP

OS Unidentified

FN JP 2002509438-A/2

PD 26-MAR-2002

PF 08-JUL-1998 JP 1999508710

PR 09-JUL-1997 US 60/053080

PI JENNY VALLADEAU, ODILE RAVEL, ELIZABETH ESTHER MARY BATES, JOHN

FORD,

PI SEM SÆLAND, SERGE J E LEBECQUE

PC C07K14/705,C12N15/12

CC Strandedness: Single;

CC Topology: Linear;

CC /note= 'short form lacks nucleotides 608-673' CC /note=

'ASGPRM (table 2) has sequence insert encoding GEE CC

between

CC nucleotides 775-776'

CC /note= 'nucleotide 1064 of DCMP2s may be

A, which would encode

CC asn rather

CC than asp at the residue numbered 270'

CC Key Location/Qualifiers

FT CDS 257..1204

FT misc feature 608

FT misc feature 775

FT misc feature 1064.

FT Location/Qualifiers

source 1..1458

		/organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"	
ORIGIN			
Query Match		100.0%; Score 1458; DB 6; Length 1458;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1458; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GTGTAGGAGATGGGATGTCCTCCAGATGATAGGGCTCTCTGGGATTTCAGACCCCAAGACCACG	60
DB	1	GTGTAGGAGATGGGATGTCCTCCAGATGATAGGGCTCTCTGGGATTTCAGACCCCAAGACCACG	60
QY	61	AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC	120
DB	61	AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC	120
QY	121	GTCCACCTCTGGTCTTGGAGACAAACCAAGCTGGGAATCAACCTCCAGACCTCCCA	180
DB	121	GTCCACCTCTGGTCTTGGAGACAAACCAAGCTGGGAATCAACCTCCAGACCTCCCA	180
QY	181	CAGCTCCACCCAGACTGGGGCCGCGCTCCATTTCCAGCTGTGACACCTCAGAG	240
DB	181	CAGCTCCACCCAGACTGGGGCCGCGCTCCATTTCCAGCTGTGACACCTCAGAG	240
QY	241	CGGTGTGGCCCAAGCATGACAAAGACGTATGAAAACCTTCCAGTACTTGGAGAAATAAGGT	300
DB	241	CGGTGTGGCCCAAGCATGACAAAGACGTATGAAAACCTTCCAGTACTTGGAGAAATAAGGT	300
QY	301	GAAAGTCCAGGGGTTTAAAAATGGGGCACTTCTCTCCAGTCCCTCTCGACGGTCTCG	360
DB	301	GAAAGTCCAGGGGTTTAAAAATGGGGCACTTCTCTCCAGTCCCTCTCGACGGTCTCG	360
QY	361	CTCTGGGCGCTGCATCTCTGCTGTCTGGGCGCTCGGCGCTGCTGCTGCTCATCAT	420
DB	361	CTCTGGGCGCTGCATCTCTGCTGTCTGGGCGCTCGGCGCTGCTGCTGCTCATCAT	420
QY	421	CTGTGTGGTGGATTCAAAATTTCCAAAATTTCCAGAGGACCTGGTGACCTCAGAACAGA	480
DB	421	CTGTGTGGTGGATTCAAAATTTCCAAAATTTCCAGAGGACCTGGTGACCTCAGAACAGA	480
QY	481	TTTTAGCAATCTCACTCAAACTGTGGGGAGATCCAGGACCTGACCTCCAGGGGAG	540
DB	481	TTTTAGCAATCTCACTCAAACTGTGGGGAGATCCAGGACCTGACCTCCAGGGGAG	540
QY	541	CAGCTTGGAGAAACCATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAACGAGGAAG	600
DB	541	CAGCTTGGAGAAACCATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAACGAGGAAG	600
QY	601	GCAGGAGGGGTATCTGAGCTCCAGGAACACACTACGCAAGAGGCACACTAGGCCACTG	660
DB	601	GCAGGAGGGGTATCTGAGCTCCAGGAACACACTACGCAAGAGGCACACTAGGCCACTG	660
QY	661	TCCCACTGCCCATCTGTGTGTGCCAGTTCATTTCTGAAATGCTCTCCGAGTCCAGCA	720
DB	661	TCCCACTGCCCATCTGTGTGTGCCAGTTCATTTCTGAAATGCTCTCCGAGTCCAGCA	720
QY	721	GCTGTGCAAGACCTGAAGAACTGACCTGCCAGGTGGCTACTCTCAACAAATGSCCTC	780
DB	721	GCTGTGCAAGACCTGAAGAACTGACCTGCCAGGTGGCTACTCTCTCAACAAATGSCCTC	780
QY	781	CATGAAGGGACCTGTGCTGCCCGCTCAACTGGGTGGAGCAACCAAGCAGCTGCTACTGGTT	840
DB	781	CATGAAGGGACCTGTGCTGCCCGCTCAACTGGGTGGAGCAACCAAGCAGCTGCTACTGGTT	840
QY	841	CTCTCACTCTGGGATGCTCTGGGCCGAGGCTGAGAAATGCTGCGAGCTGAAGAACGCCCA	900
DB	841	CTCTCACTCTGGGATGCTCTGGGCCGAGGCTGAGAAATGCTGCGAGCTGAAGAACGCCCA	900
QY	901	CCTGTGGTTCATCACTCCAGGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGC	960
DB	901	CCTGTGGTTCATCACTCCAGGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGC	960

QY	961	ATACACCTGGATGGGCTTCAGTGA	CCTCTGAAGAGGACCTTGGAGTGGGTGGATGGACAG	1020
DB	961	ATACACCTGGATGGGCTTCAGTGA	CCTCTGAAGAGGACCTTGGAGTGGGTGGATGGACAG	1020
QY	1021	CTATGGCAGCCGGCTTCCAGAA	CTGGAAAGCCAGGCCAGACGACTGGCAGGGGACACGG	1080
DB	1021	CTATGGCAGCCGGCTTCCAGAA	CTGGAAAGCCAGGCCAGACGACTGGCAGGGGACACGG	1080
QY	1081	GCTGGGTGGAGGCGGAGGACTGTGCTCACTTCCATCCAGACGGCAGGTGGAAATGACGACGT	1140	
DB	1081	GCTGGGTGGAGGCGGAGGACTGTGCTCACTTCCATCCAGACGGCAGGTGGAAATGACGACGT	1140	
QY	1141	CTGCCAGAGGCCCTTACCACTGGGTCTGCGAGGCTGGCTGGGTGAGAGAG	1200	
DB	1141	CTGCCAGAGGCCCTTACCACTGGGTCTGCGAGGCTGGCTGGGTGAGAGAG	1200	
QY	1201	TCACCTGAGCTGCTTTGGTGGGACCAACCCGGCCACAGAAATGGCGGTGGGAGGAGACTC	1260	
DB	1201	TCACCTGAGCTGCTTTGGTGGGACCAACCCGGCCACAGAAATGGCGGTGGGAGGAGACTC	1260	
QY	1261	TTCTCACGACCTCTCTCGCAAGACCGCTCTGGGAGAGAAATAAGCACCTGGGAGATTGGAAG	1320	
DB	1261	TTCTCACGACCTCTCTCGCAAGACCGCTCTGGGAGAGAAATAAGCACCTGGGAGATTGGAAG	1320	
QY	1321	CACCTGCTAACATTTTGAATTTTCTTTTAAATTTTAAAAAGATGGTATAGTGTCTTTA	1380	
DB	1321	CACCTGCTAACATTTTGAATTTTCTTTTAAATTTTAAAAAGATGGTATAGTGTCTTTA	1380	
QY	1381	AGCTTTTATTTTTCACACTTTTGAAGTCAACTTCATGCAAGGTATATTTTACATA	1440	
DB	1381	AGCTTTTATTTTTCACACTTTTGAAGTCAACTTCATGCAAGGTATATTTTACATA	1440	
QY	1441	ATAAAATGCACTCATTT	1458	
DB	1441	ATAAAATGCACTCATTT	1458	
RESULT 2				
AR561599				
LOCUS	AR561599	1458 bp	DNA	linear PAT 08-OCT-2004
DEFINITION	Sequence 3 from patent US 6756478.			
ACCESSION	AR561599			
VERSION	AR561599.1 GI:53974518			
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 1458)			
AUTHORS	Valladeau, J., Ravel, O., Bates, E.E.M., Ford, J., Saeland, S. and Lebecque, S.J.E.			
TITLE	Isolated mammalian membrane protein genes; related reagents			
JOURNAL	Patent: US 6756478-A 3 29-JUN-2004;			
FEATURES	Location/Qualifiers			
source	1..1458			
	/organism="unknown"			
	/mol_type="genomic DNA"			
ORIGIN				
Query Match 100.0%; Score 1458; DB 6; Length 1458;				
Best Local Similarity 100.0%; Pred. No. 0;				
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	GTGTAGGAGATGGGATGTCCTCCAGATGATAGGGCTCTCTGGGATTTCAGACCCAGACCCAGC	60	
DB	1	GTGTAGGAGATGGGATGTCCTCCAGATGATAGGGCTCTCTGGGATTTCAGACCCAGACCCAGC	60	
QY	61	AGGACTCCAGTCACTCTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC	120	
DB	61	AGGACTCCAGTCACTCTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC	120	
QY	121	GTCCCACTCTGGTCTCTGAGCAACCAAGCTGGGAATCAACCTCCAGACCTCCCA	180	
DB	121	GTCCCACTCTGGTCTCTGAGCAACCAAGCTGGGAATCAACCTCCAGACCTCCCA	180	

QY 181 CAGCTCCACCCAGACGTGGGGCGGGCCCTGCTCCATTTTCAGTGTGACAACTCAGAG 240
Db 181 CAGCTCCACCCAGACGTGGGGCGGGCCCTGCTCCATTTTCAGTGTGACAACTCAGAG 240
QY 241 CCCTGTTGGCCCAAGCATGACAGGAGCTATGAAACTTCCAGTACTTGGAGAAATAGGT 300
Db 241 CCCTGTTGGCCCAAGCATGACAGGAGCTATGAAACTTCCAGTACTTGGAGAAATAGGT 300
QY 301 GAAAGTCCAGGGGTTTAAATAATGGCCACTTCTCTCCAGTCCCTCTCTGAGGCTCTCG 360
Db 301 GAAAGTCCAGGGGTTTAAATAATGGCCACTTCTCTCCAGTCCCTCTCTGAGGCTCTCG 360
QY 361 CTCTGGGCGCTGCATCTCTGCTGTCTCTGCTGCTGGGCTCTGGGCTCTGCTGCTCAT 420
Db 361 CTCTGGGCGCTGCATCTCTGCTGTCTCTGCTGCTGGGCTCTGGGCTCTGCTGCTCAT 420
QY 421 CTCTGTTGTTGGATTCCAAATTTCCAAATTTTCAGAGGGAGCTGCTGACCTCTGAGAACAG 480
Db 421 CTCTGTTGTTGGATTCCAAATTTCCAAATTTTCAGAGGGAGCTGCTGACCTCTGAGAACAG 480
QY 481 TTTTAGCACTTCCACTCAACACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGCAG 540
Db 481 TTTTAGCACTTCCACTCAACACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGCAG 540
QY 541 CAGCTTGGAAAGAAACGATGATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACG 600
Db 541 CAGCTTGGAAAGAAACGATGATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACG 600
QY 601 GCAGGCGAGGGGTATCTGAGCTCCAGGAACACTAGCAGCAAGGCAACCTTAGGCCACTG 660
Db 601 GCAGGCGAGGGGTATCTGAGCTCCAGGAACACTAGCAGCAAGGCAACCTTAGGCCACTG 660
QY 661 TCCCACTGCCCCTCTGCTGTGTGCTGCCAGTTCTTCTGAAATGCTCTCGAGTCCAGCA 720
Db 661 TCCCACTGCCCCTCTGCTGTGTGCTGCCAGTTCTTCTGAAATGCTCTCGAGTCCAGCA 720
QY 721 GCTGTTGCAAGACTGAGAAACTGACCTGCGAGTGGCTTACTCTCAACAACTGCTC 780
Db 721 GCTGTTGCAAGACTGAGAAACTGACCTGCGAGTGGCTTACTCTCAACAACTGCTC 780
QY 781 CACTGAAGGCACTGCTGCCCGCTCAACTGSGTGGAGCAACAGACAGCTGCTACTGGTT 840
Db 781 CACTGAAGGCACTGCTGCCCGCTCAACTGSGTGGAGCAACAGACAGCTGCTACTGGTT 840
QY 841 CTCTCACTCTGGGATGCTCTGGGCGGAGGCTGAGAGTACTGCCAGCTGAAGAACGCCCA 900
Db 841 CTCTCACTCTGGGATGCTCTGGGCGGAGGCTGAGAGTACTGCCAGCTGAAGAACGCCCA 900
QY 901 CTTGTTGTTGTCATCACTCCAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCG 960
Db 901 CTTGTTGTTGTCATCACTCCAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCG 960
QY 961 ATACACTGATGGCTCAGTGCCTGAGGAGCCTGAGGAGCTGGTGGATGGAACAG 1020
Db 961 ATACACTGATGGCTCAGTGCCTGAGGAGCCTGAGGAGCTGGTGGATGGAACAG 1020
QY 1021 CTATGCAACCGGCTTCCAGAACTGGAAGCCAGGCGCCAGACAGCTGGCAGGCGCACGG 1080
Db 1021 CTATGCAACCGGCTTCCAGAACTGGAAGCCAGGCGCCAGACAGCTGGCAGGCGCACGG 1080
QY 1081 GCTGGTGGAGGCGAGGACTGTGCTCACTTCCATTCAGACGGCAGGTTGGAATGACGAGT 1140
Db 1081 GCTGGTGGAGGCGAGGACTGTGCTCACTTCCATTCAGACGGCAGGTTGGAATGACGAGT 1140
QY 1141 CTGCCAGAGCCCTTACCACTGGGTCTGCGAGGCTGGCTGGTCCAGACCCAGCCAGGAG 1200
Db 1141 CTGCCAGAGCCCTTACCACTGGGTCTGCGAGGCTGGCTGGTCCAGACCCAGCCAGGAG 1200
QY 1201 TCACTGAGCTGCTTTGGTGGGACCACTGGGCGCACAGAAATGGCGGTGGGAGGAGACTC 1260
Db 1201 TCACTGAGCTGCTTTGGTGGGACCACTGGGCGCACAGAAATGGCGGTGGGAGGAGACTC 1260

QY 1261 TTCTCAGACCTCTCGCAAGACCGCTCTGGGAGAGAAATAAGCACTGGGAGATTGGAAG 1320
Db 1261 TTCTCAGACCTCTCGCAAGACCGCTCTGGGAGAGAAATAAGCACTGGGAGATTGGAAG 1320
QY 1321 CACTGCTAACATTTGAAATTTTTTTCTCTTTAAATTTTAAAGATGATAGTGTCTTA 1380
Db 1321 CACTGCTAACATTTGAAATTTTTTTCTCTTTAAATTTTAAAGATGATAGTGTCTTA 1380
QY 1381 AGCTTTTATTTTTTTTCCAACTTTTGAAGTCACTTCATGAAGGTATATTTTACATA 1440
Db 1381 AGCTTTTATTTTTTTTCCAACTTTTGAAGTCACTTCATGAAGGTATATTTTACATA 1440
QY 1441 ATAAAAATGCACATTT 1458
Db 1441 ATAAAAATGCACATTT 1458

BC039011 1769 bp mRNA linear PRI 30-JUN-2004
Homo sapiens C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 14 (macrophage-derived), transcript variant 1, mRNA (cDNA clone MGC:47699 IMAGE:5751735), complete cds.

BC039011 GI:24660336
MGC.
Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shemmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Scapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalios,D.E., Schnerch,A., Schein,J.E., Jones,S.U. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 1769)
Strausberg,R.
Direct Submission
Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,

REMARK
COMMENT

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric-Pop, Legaspi, R., Maduro, Q.L., Masiello, C., Mackeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsongeon, C., Vogt, J.-L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 82 Row: j Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA q1: 54533683.

FEATURES

source

```

source
1. .1769
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="MGC:47699 IMAGE:5751735"
/tissue_type="Brain, lung, Testis, adult, pooled whole"
/clone_lib="NIH MGC_115"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

gene
1. .1769
/gene="CLECSF14"
/note="synonyms: HML2, HML, CLECSF13"
/db_xref="LocusID:10462"
/db_xref="MIM:605999"
252. .1202
/gene="CLECSF14"
/codon_start=1
/product="C-type lectin, superfamily member 13, isoform 1"
/protein_id="AAH39011.1"
/db_xref="GI:24660337"
/db_xref="LocusID:10462"
/db_xref="MIM:605999"
/translation="MTRTYENFQYLENKVKVQGFKNGLPLQSLQRLRCSCPGCHLLLS
LGLGLLAVIICVGFQNSKQRDLDTDFNSTNTVAEIQALTSQSSSLEETIA
SLKAVEGFKBERQAGSELBQHTTQKALHGCHPCSVCPVHSEMLLRVQQLVQDL
INSREQNFQVKYLGSAVTGMVDPGKAWKVDGTDYATGFGQWKPGQPDWQGGHLL
GGGECAHFHPDGRWDDVDCORPYHWVCEAGLGQTSOESH"

```

ORIGIN

[illegible]

[illegible]

Db	1180	AAGC	ATGGAGATGGAAGC	ATGCTTAACATTTTTTGAATTTTTTCTTTAAATTTTAA	1233
Qy	1361	AAAG	TGTAAGTCTTCTTAAGCTTTTATATTTTTTCCAACTTTGAAAGTCAACTTCAT	1420	
Db	1240	AAAG	TGTAAGTCTTCTTAAGCTTTTATATTTTTTCCAACTTTGAAAGTCAACTTCAT	1299	
Qy	1421	GAAG	GTATAATTTTTTACATAATAAAAAATGCACCTCATTT	1458	
Db	1300	GAAG	GTATAATTTTTTACATAATAAAAAATGCACCTCATTT	1337	
RESULT 5					
BD135983					
LOCUS	BD135983		1370 bp	DNA	linear
DEFINITION	Isolated dentritic cell membrane protein genes.				
ACCESSION	BD135983				
VERSION	BD135983.1	GI:23230928			
KEYWORDS	JP 2002509438-A/4.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	unclassified.				
AUTHORS	1 (bases 1 to 1370)				
TITLE	Valladeau,J., Ravel,O., Bates,E.E.M., Ford,J., Saeland,S. and				
JOURNAL	Lebecque,S.J.E.				
COMMENT	Isolated dentritic cell membrane protein genes				
	Patent: JP 2002509438-A 4 26-MAR-2002;				
	SCHERING CORP				
	OS	Unidentified			
	PN	JP 2002509438-A/4			
	PD	26-MAR-2002			
	PF	08-JUL-1998 JP 1999508710			
	PR	09-JUL-1997 US 60/053080			
	PI	JENNY VALLADEAU, ODILE RAVEL, ELIZABETH ESTHER MARY BATES, JOHN			
	PI	FORD,			
	PI	SEM SAE LAND, SERGE J E LEBECQUE			
	PC	C07K14/705,C12N15/12			
	CC	Strandedness: Single;			
	CC	Topology: Linear;			
	CC	Isolated dentritic cell membrane protein genes	PH	Key	
	CC	Location/Qualifiers			
FT	CDS	Location/Qualifiers	273..1091.		
FEATURES	source				
	1..1370				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	76.0%; Score 1108; DB 6; Length 1370;				
Best Local Similarity	89.4%; Pred. No. 7.9e-252;				
Matches 1288;	Conservative	0;	Mismatches 5;	Indels 147;	Gaps 3;
Qy	28	TAGGGCTCTGGGATTTTCAGACCCCAAGACGAGGAGCTCCAGTCACTTACCCAGCT	87		
Db	44	TACTGTCCCTGGGATTTTCAGATCCCAAGACCAAGCAGGAGCTCCAGTCACTTACCCAGCT	103		
Qy	88	CTCCAGGACACAGCGCTCCCAACTCTGAGTGACCTCCCACTCTGGTCTTGGCAGACAA	147		
Db	104	CTCCAGGACACAGCGCTCCCAACTCTGAGTGACCTCCCACTCTGGTCTTGGCAGACAA	163		
Qy	148	CCAACTGGGAATCACACCTCCAGACCTCCCAAGCTCCACCCAGACTGGGCGCCGGC	207		
Db	164	CCAACTGGGAATCACACCTCCAGACCTCCCAAGCTCCACCCAGACTGGGCGCCGGC	223		
Qy	208	CTCGCTCCATTTTCAGCTGTGACAACTTCAGAGCGGTGGTGGCCCAAGCATGACAAGAC	267		
Db	224	CTCGCTCCATTTTCAGCTGTGACAACTTCAGAGCGGTGGTGGCCCAAGCATGACAAGAC	283		
Qy	268	GTATGAAACCTTCCAGTACTTGGAGATAAGGTGAAAGTCCAGGGGTTTAAAAATGGCC	327		
Db	284	GTATGAAACCTTCCAGTACTTGGAGATAAGGTGAAAGTCCAGGGGTTTAAAAATGGCC	343		
Qy	328	ACTTCCCTCCAGTCCCTCTCTGCAGCGTCTCCGCTCTGGGGCCCTGCAATCTCTCTGCTGTC	387		

|||||
344 ACTTCCCTCCTCAGTCCCTC----- 362

388 CTTGGGCTCGGCCTGCTGTGCTGTGTCATCATCTGTGTGTTGGATTCCAATAATTCCAA 447

363 -----CTGCTGCTGGTCATCATCTGTGTGTTGGATTCCAATAATTCCAA 406

448 ATTTTCAGAGGGAACCTGGTGACCCTGAGAAACAGATTTTTAGCAACTTCACCTCAAAACACTGT 507

407 ATTTTCAGAGGGAACCTGGTGACCCTGAGAACAGATTTTTAGCAACTTCACCTCAAAACACTGT 466

508 GCGGAGATCCAGGCACTGACTTCCAGGGCAGCAGCTTTGGAAGAAACGATAGCATCTCT 567

467 GCGGAGATCCAGGCACTGACTTCCAGGGCAGCAGCTTTGGAAGAAACGATAGCATCTCT 526

568 GAAGAGCTGAGTGGAGGGTTTCAAGCAGGAAACGGCAGGCGAGGGTATCTGAGCTCCAGA 627

527 GAAGAGCTGAGTGGAGGGTTTCAAGCAGGAAACGGCAGG----- 564

628 ACACACTACGAGAAAGGCACACCTAGGCCAATGTCTCCCACTGCCCATCTGTGTGTGTCCC 687

565 -----C 565

688 AGTTCAATCTGAAATGCTCTCTGAGTCCAGCAGCTGGTGCAAGACCTGAAGAAACTGAC 747

566 AGTTCAATCTGAAATGCTCTCTGAGTCCAGCAGCTGGTGCAAGACCTGAAGAAACTGAC 625

748 CTGCCAGGTGGCTACTCTCAACAACAAAT-----GCCTCCACTGAAGGACCTGCTG 798

626 CTGCCAGGTGGCTACTCTCAACAACAAATGGTGAAGGAGCCTCCACTGAAGGACCTGCTG 685

799 CCCCGTCAAATGGGTGGAGCACCAAGACAGCTGCTACTGCTTCTCTCACTCTGGGATGTC 858

686 CCCCGTCAAATGGGTGGAGCACCAAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTC 745

859 CTGGCCCCAGGCTGAGAAAGTAATGCCAGCTGAAGAAACGCCCAACCTGGTGGTCAACAATC 918

746 CTGGGCCAGGCTGAGAAAGTAATGCCAGCTGAAGAAACGCCCAACCTGGTGGTCAACAATC 805

919 CAGGAGGCGCAGNAATTTTCTCCAGAAATATCTAGGCTCCGATACACCTGAGTGGGCT 978

806 CAGGAGGAGCAGAAATTTTCTCAGAAATATCTAGGCTCCGATACACCTGAGTGGGCT 865

979 CAGTGACCTGAAGAGCCTGGAAGTGGGTGGAATGGAACAGACTATGCCACCGGCTTCCA 1033

866 CAGTGAACCTGNAGGACCTGGAAGTGGGTGGATGGAACAGACTATGCCACCGGCTTCCA 925

1039 GAACTGGAAGCCAGGCCAGCCAGCAGCAGCTGGCAGGGGCAACCGGCTGGGTGGAGGCGAGA 1099

926 GAACTGGAAGCCAGGCCAGCCAGCAGCAGCTGGCAGGGGCAACCGGCTGGGTGGAGGCGAGA 985

1099 CTGTGCTCACTTCATCCAGACGGCAGTGGATGACGAGCTGTCGACAGAGCCCTTACCA 1155

986 CTGTGCTCACTTCATCCAGACGGCAGTGGAAATGACGAGCTGTCGACAGAGCCCTTACCA 1041

1159 CTGGTCTGCGAGGCTGGGCTGAGACAGCCAGGAGAGTCACTGAGCTGGCTTTCGG 1211

1046 CTGGTCTGCGAGGCTGGGCTGAGACAGCCAGGAGAGTCACTGAGCTGGCTTTCGG 1100

1219 TGGGACCAACCGGCCACAGAAAATGGCGGTGGGAGGAGGACTCTTCTCAGCACTCTCTCGC 1271

1106 TGGGACCAACCGGCCACAGAAAATGGCGGTGGGAGGAGGACTCTTCTCAGCACTCTCTCGC 1161

1279 AAGACCGCTCTGGGAGAGAAATGAGCACTGGGAGATTTGAGACACTGCTTAAATTTTGA 1333

1166 AAGACCGCTCTGGGAGAGAAATGAGCACTGGGAGATTTGAGAGCACTGCTTAAATTTTGA 1221

1339 TTTTCTCTCTTAAATTTTAAAGATGGTATAGTGTCTTAAGCTTTTATTTTTTTTCC 1399

1226 TTTTCTCTCTTAAATTTTAAAGATGGTATAGTGTCTTAAGCTTTTATTTTTTTTCC 1281

1399 AACCTTTTGAAGTCAACTTCATGAAGGTATAATTTTTTACATAATAAAAAATGCATCATTT 1451

Db	1286	AACTTTGAAAGTCAACTTCATGAGGTATAATTTTACATAATAAAAAATGCACCTCAATTT	1343
RESULT 6			
AR561601			
LOCUS	AR561601	1370 bp	DNA
DEFINITION	Sequence 9 from patent US 6756478.		linear
ACCESSION	AR561601		
VERSION	AR561601.1	GI:53974520	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1370)		
AUTHORS	Valladeau, J., Ravel, O., Bates, E. M., Ford, J., Saeland, S. and Lebecque, S. J. E.		
TITLE	Isolated mammalian membrane protein genes; related reagents		
JOURNAL	Patent: US 6756478-A 9 29-JUN-2004;		
FEATURES	Location/Qualifiers		
source	1. 1370		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		
	Query Match	76.0%;	Score 1108; DB 6; Length 1370;
	Best Local Similarity	89.4%;	Pred. No. 7.9e-252;
	Matches 1286; Conservative	0;	Mismatches 5; Indels 147; Gaps 3;
Qy	28	TAGGGCTCTCGGGATTTTCAGACCCAGACCCAGCAGGACTCCAGTCACCTCTACCCGACGT	87
Db	44	TACTGTCCCTCGGGATTTTCAGATCCAGACCCAGCAGGACTCCAGTCACCTCTACCCGACGT	103
Qy	88	CTCCAGGACACAGCGCTCCCAACTCTCAGTGTGACGTCCACCTCTGTGTCTTGCAGCACAA	147
Db	104	CTCCAGGACACAGCGCTCCCAACTCTCAGTGTGACGTCCCACTCTGTGTCTTGCAGCACAA	163
Qy	148	CCAACTGGGAATCACACCTCTCAGACTCCACAGCTCCACAGCTCCACCCAGACTGGGCGCCGC	207
Db	164	CCAACTGGGAATCACACCTCTCAGACTCCACAGCTCCACAGCTCCACCCAGACTGGGCGCCGC	223
Qy	208	CTGTGCTCCATTTACGTGTGTGACAACTCAGAGCCGTGTGGCCCAAGCATGACAAGGAC	267
Db	224	CTGTGCTCCATTTACGTGTGTGACAACTCAGAGCCGTGTGGCCCAAGCATGACAAGGAC	283
Qy	268	GTATGAAAACTTCCAGTACTTGGAGATAAGTCAAAAGTCAGGGGTTTAAAAATGGGCC	327
Db	284	GTATGAAAACTTCCAGTACTTGGAGATAAGTCAAAAGTCAGGGGTTTAAAAATGGGCC	343
Qy	328	ACTTCTCTCCAGTCCCTCTGTGACGCTCTCCGCTCTGGGCCCTGCCATCTCTCTGTCTC	387
Db	344	ACTTCTCTCCAGTCCCTC-----	362
Qy	388	CTGTGGGCTCGGCCCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	447
Db	363	-----CTGT	406
Qy	448	ATTTTCAGAGGACCTGTGTGACCTTGAGAACAGATTTTAGCAACTTCACCTCAAACTGT	507
Db	407	ATTTTCAGAGGACCTGTGTGACCTTGAGAACAGATTTTAGCAACTTCACCTCAAACTGT	466
Qy	508	GGCGGAGATCCAGGCACTGACTTCCAGGGCAGCAGCTTGGAGAAACGATAGCATCTCT	567
Db	467	GGCGGAGATCCAGGCACTGACTTCCAGGGCAGCAGCTTGGAGAAACGATAGCATCTCT	526
Qy	568	GAAGCTGAGGTGAGGGTTTCAAGCAGGAAACGGCAGGCGAGGGGTATCTGAGCTCCAGGA	627
Db	527	GAAGCTGAGGTGAGGGTTTCAAGCAGGAAACGGCAGG-----	564
Qy	628	ACACACTACGCAGAGGACACACCTTAGGCCACTGTCTCCCACTGCCCATCTGTGTGTGTGT	687
Db	565	-----	565
Qy	688	AGTTCAATTCGAAATGCTCTCGGAGTCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	747

[illegible]

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1858)

Strausberg, R.

Direct Submission

Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabps@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mc@hghri.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granice, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgion, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 49 Row: m Column: 23

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5453683.

FEATURES

Location/Qualifiers

1..1858

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:34439 IMAGE:5225381"

/tissue_type="Pancreas, Spleen, adult pooled"

/clone_lib="NIH_MGC_120"

/lab_host="DH10B"

/notes="Vector: pCMV-SPORT6"

1..1858

/genes="CLECSF14"

/note="synonyms: HML2, HML, CLECSF13"

/db_xref="LocusID:10462"

/db_xref="MIM:605999"

251..1021

/genes="CLECSF14"

/codon_start=1

/product="CLECSF14 protein"

/protein_id="AAH27858.1"

/db_xref="GI:45708678"

/db_xref="LocusID:10462"

/db_xref="MIM:605999"

/translation="MRTYVENFOYLENKVKVQKNGPUPQLQLRLRSGPCHLLLS LGGLLLVTVGVGQNSKFORDLVLTDFSNFTSNVATVIAQLTSQGSLEETIA SLKAEVEGRQERQAVHSEMLRVQQLKKLVQCVATVNNASTETGTCPPVWVE HQDSCYWFSHSGSWAEABEKYCOLKNAHLVIVINSREEQVRASGTQFLRHVPPEWVLK

gene

1..1858

/genes="CLECSF14"

/note="synonyms: HML2, HML, CLECSF13"

/db_xref="LocusID:10462"

/db_xref="MIM:605999"

251..1021

/genes="CLECSF14"

/codon_start=1

/product="CLECSF14 protein"

/protein_id="AAH27858.1"

/db_xref="GI:45708678"

/db_xref="LocusID:10462"

/db_xref="MIM:605999"

/translation="MRTYVENFOYLENKVKVQKNGPUPQLQLRLRSGPCHLLLS LGGLLLVTVGVGQNSKFORDLVLTDFSNFTSNVATVIAQLTSQGSLEETIA SLKAEVEGRQERQAVHSEMLRVQQLKKLVQCVATVNNASTETGTCPPVWVE HQDSCYWFSHSGSWAEABEKYCOLKNAHLVIVINSREEQVRASGTQFLRHVPPEWVLK

CDS

1..1858

/genes="CLECSF14"

/note="synonyms: HML2, HML, CLECSF13"

/db_xref="LocusID:10462"

/db_xref="MIM:605999"

251..1021

/genes="CLECSF14"

/codon_start=1

/product="CLECSF14 protein"

/protein_id="AAH27858.1"

/db_xref="GI:45708678"

/db_xref="LocusID:10462"

/db_xref="MIM:605999"

/translation="MRTYVENFOYLENKVKVQKNGPUPQLQLRLRSGPCHLLLS LGGLLLVTVGVGQNSKFORDLVLTDFSNFTSNVATVIAQLTSQGSLEETIA SLKAEVEGRQERQAVHSEMLRVQQLKKLVQCVATVNNASTETGTCPPVWVE HQDSCYWFSHSGSWAEABEKYCOLKNAHLVIVINSREEQVRASGTQFLRHVPPEWVLK

misc_feature		LGRTLBSGSGFQNDCHLCHTLRLDLIGLSIORNISKLLS"	
407..679		/gene="CLECSF14"	
/note="Lectin_N; Region: Hepatic lectin, N-terminal domain"			
/db_xref="CDD:pfam03954"			
710..844		/gene="CLECSF14"	
/note="CLECT; Region: C-type lectin (CTL) or carbohydrate-recognition domain (CRD)"			
/db_xref="CDD:smart00034"			
ORIGIN			
Query Match		58.8%; Score 857.4; DB 9; Length 1858;	
Best Local Similarity		73.9%; Pred. No. 2.4e-192;	
Matches 1370; Conservative		0; Mismatches 1; Indels 482; Gaps 3;	
QY	7	GAGATGGGATGTCCAGATGATAGGGCTCTGGGATTTTCAGACCCAGACGAGGACT	66
Db	1	GAGATGGGATGTCCAGATGATAGGGCTCTGGGATTTTCAGACCCAGACGAGGACT	60
QY	67	CCAGTCACCTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGACGTCCCA	126
Db	61	CCAGTCACCTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGACGTCCCA	120
QY	127	CCTCTGGTCTTTGGCAGCAACAACCGTGGGAATCACACCTCTCCAGACTCTCCACAGCTC	186
Db	121	CCTCTGGTCTTTGGCAGCAACAACCGTGGGAATCACACCTCTCCAGACTCTCCACAGCTC	180
QY	187	CACCCAGACTGGGCGCGCGCTCTCCATTTTCAGCTGTGACAACTTCAGAGCGGTGT	246
Db	181	CACCCAGACTGGGCGCGCGCTCTCCATTTTCAGCTGTGACAACTTCAGAGCGGTGT	240
QY	247	TGGCCCAAGACTGACAAGCGTATGAAACTTCACGACTTTGGAGAATAAGGTGAAAGT	306
Db	241	TGGCCCAAGACTGACAAGCGTATGAAACTTCACGACTTTGGAGAATAAGGTGAAAGT	300
QY	307	CCAGGGGTTTAAAAATGGCCCACTTCCTCTCCAGTCCCTCTCCAGCGCTCTCCGCTCTGG	366
Db	301	CCAGGGGTTTAAAAATGGCCCACTTCCTCTCCAGTCCCTCTCCAGCGCTCTCCGCTCTGG	360
QY	367	GCCTGCCATCTCTGCTGTCTGGGCTCTGGGCTCTGCTGTGCTGTGCTATCATCTGTGT	426
Db	361	GCCTGCCATCTCTGCTGTCTGGGCTCTGGGCTCTGCTGTGCTGTGCTATCATCTGTGT	420
QY	427	GTTGGATTCCAAATTTCCAAATTTCCAGAGGACCTGGTGACCTCCCTGAGAACAGATTTTAG	486
Db	421	GTTGGATTCCAAATTTCCAAATTTCCAGAGGACCTGGTGACCTCCCTGAGAACAGATTTTAG	480
QY	487	CAACTTCACCTCAAACTGTGGGGAGATCCAGSCACTGACTTCCAGGGGAGAGCTTT	546
Db	481	CAACTTCACCTCAAACTGTGGGGAGATCCAGSCACTGACTTCCAGGGGAGAGCTTT	540
QY	547	GGAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACGGCAGGC	606
Db	541	GGAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACGGCAGGC	599
QY	607	AGGGGTATCTGAGCTCCAGGAACACACTACGAGAGGACACACCTTAGGCCACTGTGCCCA	666
Db	600	-----	599
QY	667	CTGCCCATCTGTGTGTCTCCAGTTTCAATTCTGAAATGCTCTCGAGTCCAGCAGCTGGT	726
Db	600	-----CAGTTCAATCTTCAAAATGCTCTCGAGTCCAGCAGCTGGT	639
QY	727	GCAAGACCTGAAGAACTGACCTGCCAGTGGCTACTCTCAACAACATGCTCTCCACTGA	786
Db	640	GCAAGACCTGAAGAACTGACCTGCCAGTGGCTACTCTCAACAACATGCTCTCCACTGA	699
QY	787	AGGGACCTGTGCTCCCGCTCAACTGGGTGGAGCACCAAGCAGCTGTCTACTGTTCTCTCA	846
Db	700	AGGGACCTGTGCTCCCGCTCAACTGGGTGGAGCACCAAGCAGCTGTCTACTGTTCTCTCA	759

```
QY 847 CTCTGGGATCTCTGGCCGAGGCTGAGAGTACTGCCAGCTGAAGAACGCCACCTGTG 906
Db 760 CTCTGGGATCTCTGGCCGAGGCTGAGAGTACTGCCAGCTGAAGAACGCCACCTGTG 819
QY 907 GGTCAATCACTCCAGGAGGA----- 927
Db 820 GGTCAATCACTCCAGGAGGAGCAGGTGAGGGCTTCTGGTACTCAGTTCTCTAAGACATGT 879
QY 928 ----- 927
Db 880 CCCATTTAGGGAATGGTCTTAAAGCTTGGCCGACATTTGGAATCATCTGGAGCTTCCA 939
QY 928 ----- 927
Db 940 GAATGACTGTCTGTGCCACACCTTCAGAGATTTTAATAGGCTTGAGCATCCAGAGAA 999
QY 928 ----- 927
Db 1000 CATCTCCAACTCTCAGTTGATTTAAAGAGAGCTAAAGTTGAGAATCACTAGTTTAA 1059
QY 928 -----GCAGAAATTTGTCCA 942
Db 1060 GAGTTGGCTGGGATGGGAAGAGTCACTACCTGGCCCTTCTGTCAACCGCAGAAATTTGTCCA 1119
QY 943 GAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGTGACCCCTGAAGGAGCTGGAA 1002
Db 1120 GAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGTGACCCCTGAAGGAGCTGGAA 1179
QY 1003 GTGGGTGGATGGAACAGACATATGACCGCGCTTCCA----- 1038
Db 1180 GTGGGTGGATGGAACAGACATATGGGACCGCTTCCAAGTGAGTGACCCCTGGCTTCTGTG 1239
QY 1039 ----- 1038
Db 1240 CCAATGTCTCCGTACGGTTCCTCCAGGATCTGACTGGCAGCCTCTCTCCCTCGAA 1299
QY 1039 -----GACTGG 1045
Db 1300 CTGAAATCTACTCATCTACCTTCTAGCTGTGCTGTGTTCTTTTCTCTCAGGAACTGG 1359
QY 1046 AAGCAGGCCAGCAGACGACTGGCAGGGGACGGGCTGGGTGGAGGAGGAGCTGTGCT 1105
Db 1360 AAGCAGGCCAGCAGACGACTGGCAGGGGACGGGCTGGGTGGAGGAGGAGCTGTGCT 1419
QY 1106 CACTTCCATCCAGACGCGAGTGGAATGACAGCTCTGCCAGAGGCCCTTACCACTGGGTC 1165
Db 1420 CACTTCCATCCAGACGCGAGTGGAATGACAGCTCTGCCAGAGGCCCTTACCACTGGGTC 1479
QY 1166 TGGGAGGCTGGCTGGTTCAGACCCAGCAGGAGTCACTGAGCTGCCTTTGGTGGGACC 1225
Db 1480 TGGGAGGCTGGCTGGTTCAGACCCAGCAGGAGTCACTGAGCTGCCTTTGGTGGGACC 1539
QY 1226 ACCGGCCACAGAAATGGCGGTGGAGGAGGACTCTTCTCAGACCTCTCGCAAGACCG 1285
Db 1540 ACCGGCCACAGAAATGGCGGTGGAGGAGGACTCTTCTCAGACCTCTCGCAAGACCG 1599
QY 1286 CTCTGGAGAGAAATAGCACTGGAGATTGGAAGCACTGCTAACATTTTGAATTTTTT 1345
Db 1600 CTCTGGAGAGAAATAGCACTGGAGATTGGAAGCACTGCTAACATTTTGAATTTTTT 1659
QY 1346 CTCTTTAAATTTTAAAGATGGTATAGTGTCTTTAAGCTTTTATTTTTTCCACTTTT 1405
Db 1660 CTCTTTAAATTTTAAAGATGGTATAGTGTCTTTAAGCTTTTATTTTTTCCACTTTT 1719
QY 1406 GAAAGTCACTCATGAAGGTATAATTTTACATAATAAAATGCACTCATTT 1458
Db 1720 GAAAGTCACTCATGAAGGTATAATTTTACATAATAAAATGCACTCATTT 1772
```

RESULT 8

AX067341

LOCUS

DEFINITION

Sequence 45 from Patent W00078960.

661 bp

DNA

linear

PAT 24-JAN-2001

RESULT 9

RATGSLEC

LOCUS

DEFINITION

Rat Gal/GalNAc-specific lectin mRNA, complete cds.

1358 bp

mRNA

linear

ROD 27-APR-1993

AX067341
AX067341.1 GI:12544965
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Yuqiu, J. and Mitcham, J.L.
Compositions and methods for the therapy and diagnosis of breast
cancer
Patent: WO 0078960-A 45 28-DEC-2000;
CORIXA CORPORATION (US)
Location/Qualifiers
1. .661
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 41.8%; Score 606.6; DB 6; Length 661;
Best Local Similarity 98.2%; Pred. No. 7,2e-133;
Matches 612; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
QY 1 GTTGAGGAGATGGGATGCCAGATGATAGGGCTCTGGGATTTCCAGACCCAGACCCAGC 60
Db 1 GTTGAGGAGATGGGATGCCAGATGATAGGGCTCTGGGATTTCCAGACCCAGATACCAGC 60
QY 61 AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTGAC 120
Db 61 AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTGAC 120
QY 121 GTCCCACTCTGTCTTTCAGCACAAACCTGGGAATFACACCTCCAGACCTCCCA 180
Db 121 GTCCCACTCTGTCTTTCAGCACAAACCTGGGAATFACACCTCCAGACCTCCCA 180
QY 181 CAGCTCCACCCAGATGGGCGCGGCTCCCTCCATTTTCAGTGTGACAACCTCAGAG 240
Db 181 CAGCTCCACCCAGATGGGCGCGGCTCCCTCCATTTTCAGTGTGACAACCTCAGAG 240
QY 241 CCCTGTTGGCCCAAGCATGACAGGACGATGAAACTTCCAGTACTTTGGAGAATAAGT 300
Db 241 CCCTGTTGGCCCAAGCATGACAGGACGATGAAACTTCCAGTACTTTGGAGAATAAGT 300
QY 301 GAAAGTCCAGGGGTTTAAATATGGGCACTTCTCTCAGTCCCTCTCGAGGGTCTCCG 360
Db 301 GAAAGTCCAGGGGTTTAAATATGGGCACTTCTCTCAGTCCCTCTCGAGGGTCTCCG 360
QY 361 CTCTGGGCGCTGCTCTCTGCTGCTGGGCGCTGGGCGCTGCTGCTGCTGCTCATCAT 420
Db 361 CTCTGGGCGCTGCTCTCTGCTGCTGGGCGCTGGGCGCTGCTGCTGCTGCTCATCAT 420
QY 421 CTCTGTTGGTGGATTCCAAATTTCCAAATTTCCAGAGGACCTGCTGACCTTGAGAACAGA 480
Db 421 CTCTGTTGGTGGATTCCAAATTTCCAAATTTCCAGAGGACCTGCTGACCTTGAGAACAGA 480
QY 481 TTTTAGCAATTCACCTCAAACTGTGGCGAGATCCAGGCACTGACTTTCCAGGGCAG 540
Db 481 TTTTAGCAATTCACCTCAAACTGTGGCGAGATCCAGGCACTGACTTTCCAGGGCAG 540
QY 541 CAGCTTGGAGAAACCATAGCATCTCTGAAGCTGAGGTGGAGGTTTCAAGCAGGACG 600
Db 541 CAGCTTGGAGAAACCATAGCATCTCTGAAGCTGAGGTGGAGGTTTCAAGCAGGACG 600
QY 601 GCAGGAGGGGTATCTGAGCTCC 623
Db 601 GCAGGAGGGGTATCTGAAATGC 623

RESULT 9

RATGSLEC

LOCUS

DEFINITION

Rat Gal/GalNAc-specific lectin mRNA, complete cds.

1358 bp

mRNA

linear

ROD 27-APR-1993

ACCESSION J05495
VERSION GI:204302
KEYWORDS Gal/GalNac-specific lectin.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 1358)
AUTHORS Ii, M., Kurata, H., Itoh, N., Yamashina, I. and Kawasaki, T.
TITLE Molecular cloning and sequence analysis of cDNA encoding the
macrophage lectin specific for galactose and N-acetylgalactosamine
JOURNAL J. Biol. Chem. 265 (19), 11295-11298 (1990)
MEDLINE 90293078
PUBMED 2358462
COMMENT Original source text: Rat (strain Wistar) adult peritoneal
macrophage, cDNA to mRNA.
Draft entry and computer-readable sequence for [1] kindly submitted
by T. Kawasaki, 20-APR-1990, for release after publication.
FEATURES
source
1..1358
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
172..1092
/notes="Gal/GalNac-specific lectin precursor"
/codon_start=1
/protein_id="AAA41216.1"
/db_xref="GI:204303"
/translation="MTMAYENFQNLGSEKNOEAGKAPQSPFLCNILSWTHLLLSLGL
LSLLLVSVTGSQNSQRDLRTLTLDNTSNTKAEQLALASRGDSLQTLNSL
KVEDDHQGELOAQRSLQKVASLESTVEKQTLRTDLSEITDRVQQLGDKLTLC
OLASLKNNGSAVACCPPLHMEHSGSCYFWSQKWPEDAKYCOLNSNLVVNSLAE
QNFLOTHMSVVTWGLTDQNGPWRWDYTDYKEGFWHPKQPDNWFYHGLGGEDC
AFTSDGRWDDVQRPYRWCEMKLAKDS"
sig_peptide
172..195
/notes="Gal/GalNac-specific lectin signal peptide"
mat_peptide
196..1089
/product="Gal/GalNac-specific lectin"
polya_signal
1330..1335
ORIGIN
Query Match 30.0%; Score 438; DB 10; Length 1358;
Best Local Similarity 68.1%; Pred. No. 8e-93;
Matches 676; Conservative 0; Mismatches 295; Indels 21; Gaps 4;
QY 210 TGCCTTCATTTAGCTGTGACAACTCAGAGCC---GTGTTGCCCCCAAGCATGACAGGA 266
DB 122 TGTCTCGGTTTCAGTTAGTACGCTCAGAGGCTTTGTGTTGTGCAAGCATGACAAATGG 181
QY 267 CGTATCAAAACTCCAGTACTTGGAGAATAAGGTGAAAGTCCAGGGGTTTAAAAATGGGC 326
DB 182 CATATGAAAATCTCCAGAACTTGGGAGCGGAGAGAAAACCAAGAGGCTGTTAAAGC-- 239
QY 327 CACTTCTCTCCAGTCCCTCTGTCAGCGTCTCGGCTCTGGGCCCTGCCATCTCTGCTGT 386
DB 240 ---TCCTCCCAAGTCTTCTGTCGAATATCTCTCTTGGACCCACCTCTCTGTTCT 295
QY 387 CCTGGGCCCTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446
DB 296 CCTGGGCCCTTAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 355
QY 447 AATTTTCAGAGGACCTGTGTGACCTCAGAACAGATTTTAGCACTTCACTCAACACATG 506
DB 356 AGTTAAGGAGGACCTTAGAAACCTCAGAACCACTTTAGAACAAACCACTTCAACACCA 415
QY 507 TGGCGGAGATCCAGGCACCTGACTTCCAGGGCAGCAGCTTGGGAAGAAAACGATAGCATCTC 566
DB 416 AGGCTGAACCTACAGGCCCTGGCCCTCCAGGGGTGACAGCTTGCACAAACAGGAATCAATTCTC 475
QY 567 TGAAGCTGAGGTGGAGGTTTCAAGCAGGAACGGCAGGAGGGGTATCTGAGCTCCAGG 626
DB 476 TGAAGGTGGAGGTGGATGATCATGGGCAGGAACCTGACGGCAGGCCGAGGCTTGGAGCCAGA 535

QY 627 AACACACTACGAGAGGCACACACTAGGCCACTGTGCCCACTGTGCCCATCTGTGTGTGTCC 686
DB 536 AGGTGGCTTCTCTGGAGAGCACAGTGGAGAGAAGGACGACACTC-----TCAGAA 586
QY 687 CAGTTTCATTTGAAATGCTCTCGAGTGCAGAGCTGGTGCAGAGACTTGAAGAACTGA 746
DB 587 CAGATCTATCTGAAATAACCGATCGTGTGCAACAGCTGGGGAAGGACTTGAAGACCTTGA 646
QY 747 CTGCGCAGGTGGCTACTCTCAACACAACTGCTCCACTGAAGGGACCTGCGCCCGCTCA 806
DB 647 CATGTCAGCTGCCAGCCTCAAGAACAAAGCGCTCAGC---AGTGGCCTGTGCGCCCTTC 703
QY 807 ACTGGGTGGAGCACCAAGACAGCTCTACTAGTCTGCTTCTCTCACTCTGGGATGTCCTGGGCG 866
DB 704 ACTGGATGGAGCATGAAGCGAGCTGCTACTGTTCTCTCAGTCTGGGAAGCGGTGGCCTG 763
QY 867 AGGCTGAGAAGTACTGCGCAGCTGAAGAACGCCACCTCGTGGTGCATCAACTCCAGGGAGG 926
DB 764 AAGCTGACAAGTACTGCCAGCTGGAGAATTCTAACTAGTGTGTGCTCAACTCCCTGGCGG 823
QY 927 AGCAGAAATTTGTCCAGAAATATCTAGGCTCGCATACACCTGGATGGGCTCAGTGACC 986
DB 824 AGCAGAAATTTTCTACAGACTCACATGGGCACTGTGGTCACTTGGATTTGGCCTTAACGGACC 883
QY 987 CTGAAGGAGCCTGGGAAGTGGGTGGATGGAACAGACTATGCGACCGGCTTCCAGAACTGGA 1046
DB 884 AAAATGGGCCCTGGCGATGGGTGGAGCGGACCGACTATGAGAAGGCTTTACGCACTGGG 943
QY 1047 AGCCAGGCGAGCCAGACGACTGGCGAGGCGACGGGCTGGGTGGAGCGAGGACTGTGCTC 1106
DB 944 CCCCAAGCAGCCAGATAACTGTTATGACACCGGCTGGGAGGAGCGGAGGACTGTGCAC 1003
QY 1107 ACTTCATCCAGCGGAGGTGGATGACGAGCTCTGCCAGAGGCCCTACCACTGGGTCT 1166
DB 1004 ACTTCACCTCAGATGGTCTGTTGGAATGATGAGCTTTGCCAGAGGCCCTACCGCTGGGTCT 1063
QY 1167 GCGAGGCTGGCTGGGTGCTGAGCCAGCCAGGAG 1198
DB 1064 GTGAGATGAGCTGGCCAGGACAGCTAGGAG 1095
RESULT 10
AC026834
LOCUS
DEFINITION Homo sapiens chromosome 17 clone RP11-144K9 map 17, 6 unordered
pieces.
AC026834
VERSION HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 172095)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 17, clone RP11-144K9
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 172095)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Bozulavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,

Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strass, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 176813)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collimore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 13, 2002 this sequence version replaced gi:22475358.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L22894

Center clone name: 467_D_22

* NOTE: This is a 'working draft' sequence. It currently

* consists of 8 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 946: contig of 946 bp in length

* 947 1046: gap of 100 bp

* 1047 8054: contig of 7008 bp in length

* 8055 8154: gap of 100 bp

* 8155 50293: contig of 42139 bp in length

* 50294 50393: gap of 100 bp

* 50394 75289: contig of 24896 bp in length

* 75290 75389: gap of 100 bp

* 75390 98300: contig of 22911 bp in length

* 98301 98400: gap of 100 bp

* 98401 111897: contig of 13497 bp in length

* 111898 111997: gap of 100 bp

* 111998 144746: contig of 32749 bp in length

* 144747 144846: gap of 100 bp

* 144847 176813: contig of 31967 bp in length.

Location/Qualifiers

1. .176813

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosomes="17"

FEATURES
source

ORIGIN

/map="17"
/clone="RP11-467D22"
/clone_lib="RP11-467D22" Human Male BAC

Query Match	28.9%	Score 420.8	DB 2	Length 176813
Best Local Similarity	99.5%	Pred. No. 1.7e-88		
Matches 422	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY 1035	TCCAGAACTGGAAGCCAGGCCAGCCAGACGACTGCGAGGGCAGCGGGTGGAGGCG	1094		
DB 149559	TCAGAACTGGAAGCCAGGCCAGCCAGACGACTGCGAGGGCAGCGGGTGGAGGCG	149618		
QY 1095	AGGACTGTGCTCCTCCATCCAGACGCGAGGTGGAATGACGACGCTCTGCCAGAGGCCCT	1154		
DB 149619	AGGACTGTGCTCCTCCATCCAGACGCGAGGTGGAATGACGACGCTCTGCCAGAGGCCCT	149678		
QY 1155	ACCACTGGGTCTGCGAGCTGGCTGGGTCTAGACGACGAGGAGTCACTGAGCTGCCT	1214		
DB 149679	ACCACTGGGTCTGCGAGCTGGCTGGGTCTAGACGACGAGGAGTCACTGAGCTGCCT	149738		
QY 1215	TTGGTGGACCAACCCGCCAGAGAAATGCGGTGGAGGAGGACTCTTCTCAGGACCTCC	1274		
DB 149739	TTGGTGGACCAACCCGCCAGAGAAATGCGGTGGAGGAGGACTCTTCTCAGGACCTCC	149798		
QY 1275	TGCAAGAACCGCTCTGGGAGAGAAATAGCACTGGGAGATTGGAAGCACTGCTAACATTT	1334		
DB 149799	TGCAAGAACCGCTCTGGGAGAGAAATAGCACTGGGAGATTGGAAGCACTGCTAACATTT	149858		
QY 1335	TGAATTTTTTCTCTTTAAATTTTAAAGATGGTATAGTGTCTTAAAGCTTTTATTTTTT	1394		
DB 149859	TGAATTTTTTCTCTTTAAATTTTAAAGATGGTATAGTGTCTTAAAGCTTTTATTTTTT	149918		
QY 1395	TTCCAACCTTTTGAAGTCAACTTCATGAAGGTATATATTTTACATAATAAAAGTCACTC	1454		
DB 149919	TTCCAACCTTTTGAAGTCAACTTCATGAAGGTATATATTTTACATAATAAAAGTCACTC	149978		
QY 1455	ATTT 1458			
DB 149979	ATTT 149982			

RESULT 12

AC120057/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

repeat_region	complement(10678..10801)
repeat_region	/rpt_family="FLAM_C"
repeat_region	10803..10870
repeat_region	/rpt_family="MLT1A0"
repeat_region	10889..10948
repeat_region	/rpt_family="MIR"
repeat_region	complement(10949..11246)
repeat_region	/rpt_family="AluY"
repeat_region	11247..11291
repeat_region	/rpt_family="MIR"
repeat_region	complement(11305..11615)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(11616..11749)
repeat_region	/rpt_family="AluJo"
repeat_region	complement(11750..11912)
repeat_region	/rpt_family="MLT1A0"
repeat_region	complement(11962..12319)
repeat_region	/rpt_family="MLTIL"
repeat_region	complement(12434..12736)
repeat_region	/rpt_family="AluSq"
repeat_region	13661..13969
repeat_region	/rpt_family="AluY"
repeat_region	complement(13996..14307)
repeat_region	/rpt_family="AluSq"
repeat_region	complement(15041..15384)

Query Match	28.9%; Score 420.8; DB 9; Length 198821;
Best Local Similarity	99.5%; Pred. No. 1.7e-88;
Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Qy 1035	TCAGAACTGGAAAGCCAGGCGCCAGACACTGGCAGGGGCGACGGGCTGGGTGGAGGCG 1094
Db 29372	TCAGAACTGGAAAGCCAGGCGCCAGACACTGGCAGGGGCGACGGGCTGGGTGGAGGCG 29313
Qy 1095	AGGACTGTGCTCACTTCCATCCAGACGGCAGGTGGAAATGACGACGCTGCCAGAGGCCCT 1154
Db 29312	AGGACTGTGCTCACTTCCATCCAGACGGCAGGTGGAAATGACGACGCTGCCAGAGGCCCT 29253
Qy 1155	ACCACTGGGTCTGGAGGCTGGGCTGGGTCAGACACGAGGAGAGTCACTGAGCTGCCT 1214
Db 29252	ACCACTGGGTCTGGAGGCTGGGCTGGGTCAGACACGAGGAGAGTCACTGAGCTGCCT 29193
Qy 1215	TTGTGGGACCAACCCGGGCCACAGAAATGGCGGTGGGAGGAGGACTCTTCTCAGACCTCC 1274
Db 29192	TTGTGGGACCAACCCGGGCCACAGAAATGGCGGTGGGAGGAGGACTCTTCTCAGACCTCC 29133
Qy 1275	TCGCAAGACCGCTCTGGGAGAGAAATAAGCACTGGGAGATTTGAAGCAGCTGCTAACATTT 1334
Db 29132	TCGCAAGACCGCTCTGGGAGAGAAATAAGCACTGGGAGATTTGAAGCAGCTGCTAACATTT 29073
Qy 1335	TGAATTTTTTCTCTTTAAATTTTAAAAGATGGTATAGTCTCTTAAGCTTTTATTTTTT 1394
Db 29072	TGAATTTTTTCTCTTTAAATTTTAAAAGATGGTATAGTCTCTTAAGCTTTTATTTTTT 29013
Qy 1395	TTCCAACTTTTGAAGTCAACTTCATGAAGGTATAATTTTTTACATAATAAATGACATC 1454
Db 29012	TTCCAACTTTTGAAGTCAACTTCATGAAGGTATAATTTTTTACATAATAAATGACATC 28953
Qy 1455	ATTT 1458
Db 28952	ATTT 28949

RESULT 13	
AC015918/c	
LOCUS	AC015918
DEFINITION	Homo sapiens chromosome 17 clone CTD-316508 map 17, 21 unordered pieces.
ACCESSION	AC015918
VERSION	AC015918.10 GI:25140142
KEYWORDS	HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_CANCELLED.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9
10	10	10	10	10	10
11	11	11	11	11	11
12	12	12	12	12	12
13	13	13	13	13	13
14	14	14	14	14	14
15	15	15	15	15	15
16	16	16	16	16	16
17	17	17	17	17	17
18	18	18	18	18	18
19	19	19	19	19	19
20	20	20	20	20	20
21	21	21	21	21	21
22	22	22	22	22	22
23	23	23	23	23	23
24	24	24	24	24	24
25	25	25	25	25	25
26	26	26	26	26	26
27	27	27	27	27	27
28	28	28	28	28	28
29	29	29	29	29	29
30	30	30	30	30	30
31	31	31	31	31	31
32	32	32	32	32	32
33	33	33	33	33	33
34	34	34	34	34	34
35	35	35	35	35	35
36	36	36	36	36	36
37	37	37	37	37	37
38	38	38	38	38	38
39	39	39	39	39	39
40	40	40	40	40	40
41	41	41	41	41	41
42	42	42	42	42	42
43	43	43	43	43	43
44	44	44	44	44	44
45	45	45	45	45	45
46	46	46	46	46	46
47	47	47	47	47	47
48	48	48	48	48	48
49	49	49	49	49	49
50	50	50	50	50	50
51	51	51	51	51	51
52	52	52	52	52	52
53	53	53	53	53	53
54	54	54	54	54	54
55	55	55	55	55	55
56	56	56	56	56	56
57	57	57	57	57	57
58	58	58	58	58	58
59	59	59	59	59	59
60	60	60	60	60	60
61	61	61	61	61	61
62	62	62	62	62	62
63	63	63	63	63	63
64	64	64	64	64	64
65	65	65	65	65	65
66	66	66	66	66	66
67	67	67	67	67	67
68	68	68	68	68	68
69	69	69	69	69	69
70	70	70	70	70	70
71	71	71	71	71	71
72	72	72	72	72	72
73	73	73	73	73	73
74	74	74	74	74	74
75					

TITLE	JOURNAL	REFERENCE	AUTHORS
1. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
2. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
3. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
4. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
5. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
6. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
7. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
8. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
9. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
10. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1970, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Psychology	1975, Vol. 66, No. 3, pp. 1-10
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 1-10
5. The Role of the School in the Community	Journal of Educational Research	1985, Vol. 88, No. 5, pp. 1-10
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1990, Vol. 93, No. 6, pp. 1-10
7. The Importance of Teacher Professionalism	Journal of Educational Research	1995, Vol. 98, No. 7, pp. 1-10
8. The Effect of Teacher Salary on Student Achievement	Journal of Educational Research	2000, Vol. 103, No. 8, pp. 1-10
9. The Role of the Teacher in the 21st Century	Journal of Educational Research	2005, Vol. 108, No. 9, pp. 1-10
10. The Impact of Teacher Training on Student Achievement	Journal of Educational Research	2010, Vol. 113, No. 10, pp. 1-10

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 220581)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone CTD-316508
Unpublished
2 (bases 1 to 220581)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouckgalter,B., Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gaidan,J., Gaidan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karigas,A., Klein,J., Lehotzky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testkaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 220581)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,W., Johnson,K., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mlungu,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 20, 2002 this sequence version replaced gi.25046454.
All repeats were identified using RepeatMasker.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 21 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	5786: contig of 5786 bp in length
5787	5886: gap of 100 bp
5887	9996: contig of 4110 bp in length
9997	10096: gap of 100 bp
12097	12097: contig of 2808 bp in length
12905	13004: gap of 100 bp

Qy	1395	TTCAAACCTTTTGAAGTCAACTTCATGAAGGTATAATTTTTACATAATAAAAAATGCACCTC	1454
Db	98126	TTCAAACCTTTTGAAGTCAACTTCATGAAGGTATAATTTTTACATAATAAAAAATGCACCTC	9806
Qy	1455	ATTT 1458	
Db	98066	ATTT 98063	
RESULT 14			
S36676			
LOCUS	S36676	1414 bp	linear
DEFINITION	galactose and N-acetylgalactosamine-specific lectin [mice, macrophage, mRNA, 1414 nt].		
ACCESSION	S36676		
VERSION			
KEYWORDS	S36676.1	GI:249360	
SOURCE	Mus sp.		
ORGANISM	Mus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1414)		
AUTHORS	Sato,M., Kawakami,K., Osawa,T. and Toyoshima,S.		
TITLE	Molecular cloning and expression of cDNA encoding a galactose/N-acetylgalactosamine-specific lectin on mouse tumoricidal macrophages		
JOURNAL	J. Biochem. 111 (3), 331-336 (1992)		
MEDLINE	92268032		
PUBMED	1587794		
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI Gibbsq 10419] from the original journal article.		
FEATURES	Location/Qualifiers		
source	1. .1414		
	/organism="Mus sp."		
	/mol_type="mRNA"		
	/db_xref="taxon:10095"		
gene	1. .1414		
	/gene="galactose and N-acetylgalactosamine-specific lectin, MMGL"		
CDS	171..1085		
	/gene="galactose and N-acetylgalactosamine-specific lectin, MMGL"		
	/notes="Method: conceptual translation with partial peptide sequencing; MMGL"		
	/codon_start=1		
	/product="galactose and N-acetylgalactosamine-specific lectin"		
	/protein_id="AAB22171.1"		
	/db_xref="GI:249361"		
	/translations="MVYENLQNSRIEKTQPGKSPQSLRWILSWTHLLFSGLSL LLLVVVVGSGNSQLRRDLTLRATLNTTSKIAEFOSLDRADSPFKGIVSLKLV DVEDHQEIQAGRDLSQKVTLSSTVEKREQALKTDLSDLTDHVQQLRKDLKALTCQI ANLKNSEVACPLPHTHEGSCYWFSEKSPADKYCRLENSHLVVVNSLEQGN FLQRLANVYVSWTGLTDQNGPWRVVDGTDPEKGFKNWAPLQPDNWFHGLGGEDCRAH IITGPGWNDDVCQRTFRWICEMKLAKES"		
ORIGIN			
Query Match	27.8%;	Score 404.6;	DB 10; Length 1414;
Best Local Similarity	66.4%;	Prod. No. 6.8e-85;	
Matches	635; Conservative	0; Mismatches 304; Indels 18; Gaps 3;	
Qy	242	CGTGTTCGGCCACGACATGACAAGGACCTATGAAAACCTTCACGTACTTTGGAGATAAGGTG	301
Db	150	CGTGTTCGTTTCAGTGTAAACATGATATACGAAAACCTCCAGAACTCAAGATCGAGAG	209
Qy	302	AAAGTCACGAGGGTTTAAAAATGGGCCACTTCCTCTCTCCAGTCCCTCTCTGCAGCGTCTCCGC	361
Db	210	AAACCCAAAGAGCTGTGTAAGC-----TCCITCCAGTCCCTCTCTGTGGCGATCCTC	263
Qy	362	TCTGGGCGCTGCCATCTCTCTGCTGTCTCTGGGCGCTCGGCCCTGCTGCTGCTGCTCATCATC	421
Db	264	TCTTGACCCACCTCTCTCTGTTCTCTCCCTGGGCGCTCAGCCTCTGCTGTTGGTGGTGGTGC	323

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 19:09:15 ; Search time 838 Seconds
(without alignments)
10299.502 Million cell updates/sec

Title: US-10-829-107-3

Perfect score: 1458

Sequence: 1 gttgagagatgggatgtcc.....taataaaatgcactcattt 1458

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1458	100.0	1458	2	AAx04866
2	1434.2	98.4	1531	6	ABK94927
3	1108	76.0	1370	2	AAx04868
4	1039.2	71.3	1072	8	ACA10155
5	724.2	49.7	929	3	AAf17999
6	717.6	49.2	944	12	ADO08342
7	714.6	49.0	820	8	ACA10154
8	714.6	49.0	820	12	ADO08340
9	606.6	41.6	661	4	AAf44889
10	438	30.0	1358	10	ABT42075
11	434	29.8	1432	13	ACN41226
12	431.4	29.6	1477	13	ACN41224
13	382	26.2	1277	6	ABN95732
14	382	26.2	1277	10	ACA56588
15	382	26.2	1277	12	ADI56384
16	380.4	26.1	1277	2	AAT66950
17	380.4	26.1	1295	4	AAH57466
18	336.6	23.1	855	10	ADB52510
19	321.4	22.0	1055	9	ACD06173
20	321.4	22.0	1300	2	AAT66951

21	321.4	22.0	1300	6	ABN97236	Abn97236 Gene #373
22	313.4	21.5	1055	12	ADN95955	Adn95955 Human NOV
23	312.6	21.4	1276	13	ACN41228	Acn41228 Human dia
24	310.4	21.3	361	3	AAC09703	Aac09703 Human sec
25	307.6	21.1	1418	13	ACN41227	Acn41227 Human dia
26	305	20.9	1463	13	ACN41225	Acn41225 Human dia
27	303	20.8	1112	9	ACD06172	AcD06172 Human CDN
28	299.8	20.6	1112	12	ADN95953	Adn95953 Human NOV
29	282.8	19.4	1309	6	ABN95737	Abn95737 Gene #223
30	282.8	19.4	1309	9	ACD06171	AcD06171 Human CDN
31	282.8	19.4	1309	10	ADN84950	Adn84950 Farnesyl
32	282.8	19.4	1493	13	ADP23958	Adp23958 PRO polyP
33	267.2	18.3	1309	12	ADN95951	Adn95951 Human NOV
34	262.4	18.0	780	12	ADO21128	Ado21128 Human car
35	246	16.9	443	8	ABX41898	Abx41898 Bovine ES
36	233	16.0	1346	13	ACN41223	Acn41223 Human dia
37	233	16.0	1391	13	ACN41222	Acn41222 Human dia
38	231.8	15.9	466	9	ACH21572	Ach21572 Human adu
39	220	15.1	566	4	AAH57258	Aah57258 Human liv
40	212.4	14.6	1290	6	ABK63735	Abk63735 Rat seque
41	212.4	14.6	1290	10	ADB58277	Adb58277 Toxicity-
42	212.4	14.6	1290	10	ADB582814	Adb582814 Primary r
43	211	14.5	521	13	ADQ58274	Adq58274 Novel can
44	183	12.6	400	3	AAC57412	Aac57412 Arachidon
45	155.6	10.7	448	3	AAC57415	Aac57415 Arachidon

ALIGNMENTS

RESULT 1
AAx04866
ID AAX04866 standard; CDNA; 1458 BP.
XX
AC AAX04866;
XX
DT 11-MAY-1999 (first entry)
XX
DE Primate DCM2 C-lectin family gene nucleotide sequence.
XX
KW ss; Primate; dendritic cell membrane protein; DCM2; DCM2;
KW chromosomal abnormality; expression misregulation;
KW abnormal proliferation; regeneration; degeneration; haematopoietic cell.
XX
OS Mammalia.
XX
PH Key Location/Qualifiers
FT CDS 257..1207
FT /*tag= a
FT /product= "DCMP2 C-lectin family gene protein"
XX
XX WO9902562-A1.
XX
PD 21-JAN-1999.
XX
PF 08-JUL-1998; 98WO-US013436.
XX
PR 09-JUL-1997; 97US-0053080P.
XX
PA (SCHE) SCHERING CORP.
XX
XX Valladeau J, Ravel O, Bates EEM, Ford J, Saeland S, Lebecque SJE;
XX WPI; 1999-120786/10.
XX
PT Dendritic cell membrane proteins - used to treat conditions associated
XX with abnormal physiology or development.
XX
PS Disclosure; Page 68-69; 82pp; English.
XX
CC Dendritic cell membrane protein 1 (DCMP1) and DCM2 nucleic acids can be
XX used as markers for distinguishing cell types, including genomic aspects
XX of cells, as well as mRNA and protein expression patterns. They can also

CC be used to detect chromosomal abnormalities. The proteins can be used to
CC diagnose disorders associated with expression misregulation. They can
CC also be used to treat conditions associated with abnormal physiology or
CC development, including abnormal proliferation, e.g. cancerous conditions
CC or degenerative conditions. Abnormal proliferation, regeneration, the
CC degeneration and atrophy may be modulated using the proteins. The
CC proteins may also play a role in regulation or development of
CC haematopoietic cells
XX
SQ

Sequence 1458 BP; 349 A; 406 C; 388 G; 315 T; 0 U; 0 Other;

Query Match 100.0%; Score 1458; DB 2; Length 1458;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTGAGGAGATGGGATGCTCCAGATCATAGGGCTCTGGGATTTGAGCCCAAGCAGC 60
DB 1 GTTGAGGAGATGGGATGCTCCAGATCATAGGGCTCTGGGATTTGAGCCCAAGCAGC 60
QY 61 AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC 120
DB 61 AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC 120
QY 121 GTCCCACTCTGTGCTTTCAGCACAAACAAAGTGGGAATCAACCTCCAGACTCTCCCA 180
DB 121 GTCCCACTCTGTGCTTTCAGCACAAACAAAGTGGGAATCAACCTCCAGACTCTCCCA 180
QY 181 CAGCTCCACCCAGATGGGCGCGCGCTCTCCAGTCTCACTTTCAGCTGTGACCACTCAGAG 240
DB 181 CAGCTCCACCCAGATGGGCGCGCGCTCTCCAGTCTCACTTTCAGCTGTGACCACTCAGAG 240
QY 241 CCGTGTGGCCCAAGCATGACAGGACGTATGAAAATTTCCAGTACTTGGAGAAATAGGT 300
DB 241 CCGTGTGGCCCAAGCATGACAGGACGTATGAAAATTTCCAGTACTTGGAGAAATAGGT 300
QY 301 GAAAGTCCAGGGGTTTAAATAATGGGCCACTTCTCTCCAGTCCCTCTCCAGCGTCTCCG 360
DB 301 GAAAGTCCAGGGGTTTAAATAATGGGCCACTTCTCTCCAGTCCCTCTCCAGCGTCTCCG 360
QY 361 CTCTGGGCGCTGCACTCTCTGCTGTCTCTGGGCTCTGGGCTCTGGCTGTGTCATCAT 420
DB 361 CTCTGGGCGCTGCACTCTCTGCTGTCTCTGGGCTCTGGGCTCTGGCTGTGTCATCAT 420
QY 421 CTGTGTGGTTGGATTCAAAATTTCCAAAATTTCCAGAGGACCTGGTGACCTCAGAACAGA 480
DB 421 CTGTGTGGTTGGATTCAAAATTTCCAAAATTTCCAGAGGACCTGGTGACCTCAGAACAGA 480
QY 481 TTTTAGCAACTTCACTCAAAACACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGCAG 540
DB 481 TTTTAGCAACTTCACTCAAAACACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGCAG 540
QY 541 CAGCTTGGAGAACATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACG 600
DB 541 CAGCTTGGAGAACATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACG 600
QY 601 GCAGGACAGGGTATCTGAGCTCCAGAAACACACTAGCAGAGAGGACACCTAGGCCACTG 660
DB 601 GCAGGACAGGGTATCTGAGCTCCAGAAACACACTAGCAGAGAGGACACCTAGGCCACTG 660
QY 661 TCCCACTGCCACTGTGTGTGCCAGTTCATTCTGAAATGCTCCTCGGAGTCCAGCA 720
DB 661 TCCCACTGCCACTGTGTGTGCCAGTTCATTCTGAAATGCTCCTCGGAGTCCAGCA 720
QY 721 GCTGTGTCAGACCTGAAGAACTGACCTGCCAGGTGGCTACTCTCAACAAATGCTCTC 780
DB 721 GCTGTGTCAGACCTGAAGAACTGACCTGCCAGGTGGCTACTCTCAACAAATGCTCTC 780
QY 781 CACTGAAGGACCTGTGCTGCCCTCAACTGGGTGGAGACCAAGACAGCTGTACTGTGTT 840
DB 781 CACTGAAGGACCTGTGCTGCCCTCAACTGGGTGGAGACCAAGACAGCTGTACTGTGTT 840
QY 841 CTCTCACTCTGGGATGTCTTGGGCGCAGGCTGAGAAATGTAAGTCTGCACTGAAAGACGCCCA 900
DB 841 CTCTCACTCTGGGATGTCTTGGGCGCAGGCTGAGAAATGTAAGTCTGCACTGAAAGACGCCCA 900

DB 841 CTCTCACTCTGGGATGTCTTGGGCGCAGGCTCAGAAAGTACTGCCAGCTGAAAGAACGCCCA 900
QY 901 CCTGTGTGTCATCACTCACTCCAGGAGGAGCAGAAATTTGTCCAGAAATATCTAGGCTCCGC 960
DB 901 CCTGTGTGTCATCACTCACTCCAGGAGGAGCAGAAATTTGTCCAGAAATATCTAGGCTCCGC 960
QY 961 ATACACCTTGGATGGGCTTCACTGAGCCCTGAAAGGACCTGGAAGTGGGTGGATGGAACAGA 1020
DB 961 ATACACCTTGGATGGGCTTCACTGAGCCCTGAAAGGACCTGGAAGTGGGTGGATGGAACAGA 1020
QY 1021 CTATGCGACCGGCTTCCAGAACTGGAAGCCAGGCCAGCCAGCAGCTGGCAGGGGCACGG 1080
DB 1021 CTATGCGACCGGCTTCCAGAACTGGAAGCCAGGCCAGCCAGCAGCTGGCAGGGGCACGG 1080
QY 1081 GCTGGGTGGAGCGGAGGACTGTGCTCACTTCCATCCAGACCGCAGGTGGAATGACGAGT 1140
DB 1081 GCTGGGTGGAGCGGAGGACTGTGCTCACTTCCATCCAGACCGCAGGTGGAATGACGAGT 1140
QY 1141 CTGCCAGAGGCGCTTACCACTGGGTCTGCGAGGCTGGCTGGGTCTGAGACCCAGCCAGGAG 1200
DB 1141 CTGCCAGAGGCGCTTACCACTGGGTCTGCGAGGCTGGCTGGGTCTGAGACCCAGCCAGGAG 1200
QY 1201 TCATGAGCTGCTTGGTGGGACCAACCCGGCCACAGAAATGGCGGTGGAGGAGGACTC 1260
DB 1201 TCATGAGCTGCTTGGTGGGACCAACCCGGCCACAGAAATGGCGGTGGAGGAGGACTC 1260
QY 1261 TTCTCAGCAGCTCTCGCAAGACCGCTCTGGGAGAGAAATAGCACTGGGAGATGGAG 1320
DB 1261 TTCTCAGCAGCTCTCGCAAGACCGCTCTGGGAGAGAAATAGCACTGGGAGATGGAG 1320
QY 1321 CACTCTAACATTTTGAATTTTCTTTAAATTTTAAAGATGGTATAGTGTCTTA 1380
DB 1321 CACTCTAACATTTTGAATTTTCTTTAAATTTTAAAGATGGTATAGTGTCTTA 1380
QY 1381 AGCTTTATTTTTCCTTCAACTTTTGAAGTCACTTCATGAAAGTATAATTTTACATA 1440
DB 1381 AGCTTTATTTTTCCTTCAACTTTTGAAGTCACTTCATGAAAGTATAATTTTACATA 1440
QY 1441 ATAAAGTGCACCTCATTT 1458
DB 1441 ATAAAGTGCACCTCATTT 1458

RESULT 2

ABK94927
ID ABK94927 standard; cDNA; 1531 BP.

XX
AC ABK94927;

XX
DT 30-AUG-2002 (first entry)

XX
TX Human novel polynucleotide #38.

DE Human; gene; ss; inflammatory condition; shock; sepsis; immune response;
XX cancer; wound healing; central nervous system disease; haematopoiesis;
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
KW bone degenerative disorder; periodontal disease; reperfusion injury;
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
KW allergic condition; thrombolytic; thrombosis; coagulation disorder;
KW fungal infection.

OS Homo sapiens.

XX WO200244340-A2.

XX 06-JUN-2002.

XX 30-NOV-2001; 2001WO-US047004.

XX 30-NOV-2000; 2000US-00728952.

XX

[illegible]

QY 1339 TTTTCTCTCTTAATTTTAAAGAGATGATAGTGTCTTAAGCTTTTATTTTTC 1398
 Db 1226 TTTTCTCTCTTAATTTTAAAGAGATGATAGTGTCTTAAGCTTTTATTTTTC 1285
 QY 1399 AACTTTTGAAGTCAACTTCATGAAGGTATAATTTTACATAATAAATGCACTATT 1458
 Db 1286 AACTTTTGAAGTCAACTTCATGAAGGTATAATTTTACATAATAAATGCACTATT 1345

RESULT 4

ACA10155
 ID ACA10155 standard; cDNA; 1072 BP.

AC ACA10155;

XX 02-JUN-2003 (first entry)

XX Human NOVX polynucleotide #45.

XX Human; NOVX; gene; ss; metabolic disorder; diabetes; infectious disease;
 obesity; anorexia; cancer; cardiovascular disorder; asthma; neurogenesis;
 neurodegenerative disorder; epilepsy; immune disorder; osteoarthritis;
 haematopoietic disorder; inflammatory skin disorder; dyslipidemia;
 haematopoiesis; wound healing; angiogenesis; bacterial infection;
 viral infection; fungal infection; helminthic infection; atherosclerosis;
 protozoal infection; hypertension.

XX Homo sapiens.

OS XX

XX W0200290504-A2.

XX PN

XX PD

XX 14-NOV-2002.

XX 02-MAY-2002; 2002WO-US014342.

XX 03-MAY-2001; 2001US-0288395P.

PR 04-MAY-2001; 2001US-0288900P.

PR 07-MAY-2001; 2001US-0289087P.

PR 14-MAY-2001; 2001US-0290753P.

PR 15-MAY-2001; 2001US-0291189P.

PR 16-MAY-2001; 2001US-0291243P.

PR 18-MAY-2001; 2001US-0292001P.

PR 21-MAY-2001; 2001US-0292374P.

PR 22-MAY-2001; 2001US-0292587P.

PR 23-MAY-2001; 2001US-0293107P.

PR 28-MAY-2001; 2001US-0294110P.

PR 30-MAY-2001; 2001US-0294434P.

PR 31-MAY-2001; 2001US-0294827P.

PR 18-JUN-2001; 2001US-0298988P.

PR 31-JUL-2001; 2001US-0308901P.

PR 17-AUG-2001; 2001US-0313388P.

PR 21-AUG-2001; 2001US-0313851P.

PR 21-AUG-2001; 2001US-0313937P.

PR 17-SEP-2001; 2001US-0322701P.

PR 17-SEP-2001; 2001US-0322802P.

PR 25-SEP-2001; 2001US-0324757P.

PR 27-SEP-2001; 2001US-0325314P.

PR 27-SEP-2001; 2001US-0325682P.

PR 21-NOV-2001; 2001US-0332129P.

PR 03-DEC-2001; 2001US-0336882P.

PR 14-DEC-2001; 2001US-0340305P.

PR 01-MAY-2002; 2002US-00138588.

XX (CURA-) CURAGEN CORP.

XX Alsobrook JP, Anderson DM, Boldog FL, Burgess CE, Casman SJ;

PI Chapoval A, Edinger S, Gerlach V, Gorman L, Gunther E, Guo X;

PI Kekuda R, Lepley DM, Li L, Liu X, Malyankar UM, Miller CE;

PI Millet I, Padigaru M, Patturajan M, Pena CEA, Rieger DK, Shenoy SG;

PI Shimkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ;

PI Zerhusen BD;

XX WPI; 2003-103512/09.

DR P-PSDB; ABUS9170.

XX New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.

PS Claim 20; Page 169-170; 340pp; English.

XX The invention relates to human NOVX polypeptides and the polynucleotides
 CC encoding them. The polypeptides, polynucleotides and antibodies that bind
 CC immunospecifically to the polypeptides are useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease,
 CC preferably a NOVX-associated disorder. The sequences are useful for
 CC treating, preventing or diagnosing diseases such as metabolic disorders,
 CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,
 CC helminthic, and protozoal), anorexia, cancer, cardiovascular disorders
 CC (e.g. hypertension, atherosclerosis), neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease), epilepsy, immune disorders,
 CC osteoarthritis, haematopoietic disorders, inflammatory skin disorders,
 CC asthma and various dyslipidemias. The nucleic acids and polypeptides may
 CC also be used as targets for the identification of small molecules that
 CC modulate or inhibit e.g. neurogenesis, cell differentiation, cell
 CC proliferation, haematopoiesis, wound healing and angiogenesis, and in the
 CC generation of antibodies that bind immunospecifically to NOVX substances
 CC for use in therapeutic or diagnostic methods. The nucleic acids are
 CC further used as hybridisation probes, and in chromosome mapping, tissue
 CC typing, preventive medicine and pharmacogenomics. Sequences ACA1011-
 CC ACA10156 represent human NOVX polynucleotides of the invention
 XX
 SQ Sequence 1072 BP; 247 A; 297 C; 313 G; 215 T; 0 U; 0 Other;

Query Match 71.3%; Score 1039.2; DB 8; Length 1072;

Best Local Similarity 98.9%; Pred. No. 7.5e-259;

Matches 1060; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

QY 213 CTCATTTTCAGCTGTGACACCTCAGAGCCGTGTGCCCCAAGCATCAGAGGAGTATG 272
 Db 1 CTCATTTTCAGCTGTGACACCTCAGAGCCGTGTGCCCCAAGCATCAGAGGAGTATG 60
 QY 273 AAAAATTCAGTACTTGGAGATAAGGTGAAGTCCAGGGGTTTAAAAATGGGCCACTTC 332
 Db 61 AAAATTCAGTACTTGGAGATAAGGTGAAGTCCAGGGGTTTAAAAATGGGCCACTTC 120
 QY 333 CTCCTCAGTCCCTCTGCGAGCGTCTCGCTCTGGGCCCTGCCATCTCCTGCTCCCTGG 392
 Db 121 CTCCTCAGTCCCTCTGCGAGCGTCTCGCTCTGGGCCCTGCCATCTCCTGCTCCCTGG 180
 QY 393 GCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 452
 Db 181 GCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 QY 453 AGAGGGACCTGGTGACCCCTGAGAACAGATTTTATAGCACTTTCACCTCAAACTCTGGCGG 512
 Db 241 AGAGGGACCTGGTGACCCCTGAGAACAGATTTTATAGCACTTTCACCTCAAACTCTGGCGG 300
 QY 513 AGATCCAGGCACTGACTTCCAGGGCAGCAGCTTGGAGAAACGATAGCATCTCTGAAAG 572
 Db 301 AGATCCAGGCACTGACTTCCAGGGCAGCAGCTTGGAGAAACGATAGCATCTCTGAAAG 360
 QY 573 CTGAGTGGAGGGTTTCAACGAGGAACCGGAGGGGTATCTGAGCTCCAGGAACACA 632
 Db 361 CTGAGTGGAGGGTTTCAACGAGGAACCGGAGGGGTATCTGAGCTCCAGGAACACA 420
 QY 633 CTAGCGAAGAGGCACACCTAGGCACTGTCCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTG 692
 Db 421 CTAGCGAAGAGGCACACCTAGGCACTGTCCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 693 ATTCTGAAATGCTCTGCGAGTCCAGCAGCTGGTGCAGAACCTGGAAGAACTGACCTGCC 752
 Db 481 ATTCTGAAATGCTCTGCGAGTCCAGCAGCTGGTGCAGAACCTGGAAGAACTGACCTGCC 540
 QY 753 AGGTGGCTACTCTCAACAACAAT-----GCCTCCACTGAAGGGACCTGCTGCCCG 803

Db 541 AGGTGGCTACTCTCAACAACAAATGGTGAAGAGCCTCCACTGAAGGGACCTGCTGCCCTG 600
QY 804 TCAACTGGGTGGAGCACCAGACAGCTGTACTGGTTCTCTCACTCTGGATGTCCTGG 863
Db 601 TCAACTGGGTGGAGCACCAGACAGCTGTACTGGTTCTCTCACTCTGGATGTCCTGG 660
QY 864 CCAGAGCTGAGAAAGTACTGCGAGCTGAAGAACCCACCTGGTGGTCAATCAACTCCAGGG 923
Db 661 CCGAGGCTGGAAGTACTGCCAGCTGAAGAACCCACCTGGTGGTCAATCAACTCCAGGG 720
QY 924 AGGAGCAGAAATTTGTCGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGTG 983
Db 721 AGGAGCAGAAATTTGTCGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGTG 780
QY 984 ACCCTGAAGGAGCTGGAGTGGTGGTGAAGAACCCACCTGGTGGTCAATCAACTCCAGAACT 1043
Db 781 ACCCTGAAGGAGCTGGAGTGGTGGTGAAGAACCCACCTGGTGGTCAATCAACTCCAGAACT 840
QY 1044 GGAAGCCAGGCCAGCAGACCACTGCGAGGGGCACGGGCTGGGTGGAGCGAGGACTGTG 1103
Db 841 GGAAGCCAGGCCAGCAGACCACTGCGAGGGGCACGGGCTGGGTGGAGCGAGGACTGTG 900
QY 1104 CTCACCTTCCATCCAGACGGCAGGTGGAATGAACAGCTCTGCCAGAGGCCCTACCACTGGG 1163
Db 901 CTCACCTTCCATCCAGACGGCAGGTGGAATGAACAGCTCTGCCAGAGGCCCTACCACTGGG 960
QY 1164 TCTGCGAGGCTGGCTGGCTGACAGCAGCCAGAGAGTCACTGAGCTGCTTTGGTGGGA 1223
Db 961 TCTGCGAGGCTGGCTGGCTGACAGCAGCCAGAGAGTCACTGAGCTGCTTTGGTGGGA 1020
QY 1224 CCACCCGGCCACAGAAATGGCGTGGGAGGAGTCTTCTCACGACCTCCT 1275
Db 1021 CCACCCGGCCACAGAAATGGCGTGGGAGGAGTCTTCTCACGACCTCCT 1072

RESULT 5
ID AAF17999 standard; DNA; 929 BP.
XX AAF17999;
AC AAF17999;
XX
DT 14-MAR-2001 (first entry)
XX
DE Lung cancer associated polynucleotide sequence SEQ ID 18.
XX
KW Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulnary;
KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease; ds.
XX
OS Homo sapiens.
XX
PN WO200055180-A2.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005918.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Ruben SM;
XX
DR WPI; 2000-587514/55.
DR P-PSDB; AAB58123.
XX
PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.

XX
PS
XX
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC immunomodulatory; muscular active general; vulnary; gastrointestinal;
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences
XX
SQ Sequence 929 BP; 217 A; 267 C; 234 G; 201 T; 0 U; 10 Other;
Query Match 49.7%; Score 724.2; DB 3; Length 929;
Best Local Similarity 89.8%; Pred. No. 3.6e-177;
Matches 836; Conservative 6; Mismatches 6; Indels 83; Gaps 3;
QY 1 GTTGAGGAGATGGGATGCCAGATGATAGGCTCTCTGGATTTTCAGACCCAGACCAGC 60
Db 32 GTTGAGGAGATGGGATGCCAGATGATAGGCTCTCTGGATTTTCAGACCCAGACCAGC 91
QY 61 AGGACTCCAGTCACCTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC 120
Db 92 AGGACTCCAGTCACCTTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC 151
QY 121 GTCCACCTCTGGTCTTGGAGCAACAACGTTGGGAATCAACCTCCAGACCTCCCA 180
Db 152 GTCCACCTCTGGTCTTGGAGCAACAACGTTGGGAATCAACCTCCAGACCTCCCA 211
QY 181 CAGCTCCACCCAGATGGGCGCGGCTGCTCCATTTTCAGCTGTGACACCTCAGAG 240
Db 212 CAGCTCCACCCAGATGGGCGCGGCTGCTCCATTTTCAGCTGTGACACCTCAGAG 271
QY 241 CCGTGTGGCCCAAGCATGACAAAGACGCTATGAAATTTCCAGTACTTGGAGAAAGT 300
Db 272 CCGTGTGGCCCAAGCATGACAAAGACGCTATGAAATTTCCAGTACTTGGAGAAAGT 331
QY 301 GAAAGTCCAGGGGTTTAAAAATGGGCCACTTCTCTCCAGTCCCTCTCGAGGCTCCG 360
Db 332 GAAAGTCCAGGGGTTTAAAAATGGGCCACTTCTCTCCAGTCCCTCTCGAGGCTCCG 391
QY 361 CTCTGGGCGCTGCCATCTCTGCTGCTGGGCTCGGCTGCTGCTGCTGCTGCTCATCAT 420
Db 392 CTCTGGGCGCTGCCATCTCTGCTGCTGGGCTCGGCTGCTGCTGCTGCTCATCAT 451
QY 421 CTGTGTGGTTGGATTTCCAAATTTCCAGAGGACCTGTGGTACCTCCAGAACAGA 480
Db 452 CTGTGTGGTTGGATTTCCAAATTTCCAGAGGACCTGTGGTACCTCCAGAACAGA 511
QY 481 TTTTAGCAACTTTCACCTCAAAACACTGTGGCGAGATCCAGGCACTGACTTCCAGGGCAG 540
Db 512 TTTTAGCAACTTTCACCTCAAAACACTGTGGCGAGATCCAGGCA-TCGACTTCCAGGGCAG 570
QY 541 CAGCTTGGAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGGTTTCAACGAGGAACG 600
Db 571 CAGCTTGGAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGGTTTCAACGAGGAACG 630
QY 601 GCAGGCGAGGGGTATCTGAGCTCCAGGAAACACACTACGAGAGGACACACTAGGCCACTG 660
Db 631 GCAGG----- 635
QY 661 TCCCCACTGCCCATCTGTGTGTGTCCAGTTCAATTTGAAATGCTCTCGGAGTCCAGCA 720
|||||

```

Db 636 -----CAGTTTCATTCTGAATGCTCTCGAGTCCAGCA 669
QY 721 GCTGTCAGACCTGAAGAACTGACCTGCGAGTGGCTACTCTTCAACAACAT-GCCT 779
Db 670 GCTGGTCAAGACCTGAAGAACTGACCTGCGAGTGGCTACTCTTCAACAACATGCGCT 729
QY 780 CCACTGAAGGACCTGCTGCCCGTCAACTGGTGGAGCACCAGACAGCTGCTACTGGT 839
Db 730 CCACTGAAGGACCTGCTGCCCGTCAACTGGTGGAGCACCAGACAGCTGCTACTGGT 789
QY 840 TCTCTCACTGGGATGCTCTGGCCGAGCTGAGAACTACTGCCAGCTGAAGAACGCC 899
Db 790 TCTCTCACTGGGATGCTCTGGCCGAGCTGAGAACTACTGCCAGCTGAAGAACGCC 849
QY 900 ACTGGTGGTCATCACTCCAGGAGGAGCA 930
Db 850 ACTGGTGGTCATCACTCCAGGAGGAGCA 880

RESULT 6
AD008342
ID AD008342 standard; cDNA; 944 BP.
XX AC
XX AD008342;
XX DT
XX 01-JUL-2004 (first entry)
XX DE
XX Human NOVX polynucleotide #45.
XX Human; NOVX; gene; ss; cardiomyopathy; atherosclerosis; hypertension;
KW scleroderma; obesity; cancer; diabetes; haemophilia;
KW graft-versus-host disease; AIDS; asthma; Crohn's disease;
KW multiple sclerosis; infection; anorexia; cancer-associated cachexia;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW haematopoietic disorder; wasting disorder.
XX OS
XX Homo sapiens.
XX FN
XX US2004018594-A1.
XX PD
XX 29-JAN-2004.
XX PF
XX 01-MAY-2002; 2002US-00138588.
XX PR
XX 03-MAY-2001; 2001US-0288395P.
XX PR
XX 04-MAY-2001; 2001US-0288900P.
XX PR
XX 07-MAY-2001; 2001US-0289087P.
XX PR
XX 14-MAY-2001; 2001US-0290753P.
XX PR
XX 15-MAY-2001; 2001US-0291189P.
XX PR
XX 16-MAY-2001; 2001US-0291243P.
XX PR
XX 21-MAY-2001; 2001US-0292001P.
XX PR
XX 22-MAY-2001; 2001US-0292374P.
XX PR
XX 23-MAY-2001; 2001US-0293107P.
XX PR
XX 30-MAY-2001; 2001US-0294434P.
XX PR
XX 31-MAY-2001; 2001US-0294827P.
XX PR
XX 17-AUG-2001; 2001US-0308901P.
XX PR
XX 21-AUG-2001; 2001US-0313851P.
XX PR
XX 21-AUG-2001; 2001US-0313937P.
XX PR
XX 17-SEP-2001; 2001US-0322701P.
XX PR
XX 17-SEP-2001; 2001US-0322802P.
XX PR
XX 25-SEP-2001; 2001US-0324757P.
XX PR
XX 27-SEP-2001; 2001US-0325314P.
XX PR
XX 27-SEP-2001; 2001US-0325682P.
XX PR
XX 21-NOV-2001; 2001US-0332129P.
XX PR
XX 20-DEC-2001; 2001US-0336882P.
XX PR
XX 14-DEC-2001; 2001US-0340305P.
XX (ALSO/) ALSOBROOK J P.
PA (ANDE/) ANDERSON D W.
PA (BOLD/) BOLDOG F L.

```

```

PA (BURG/) BURGESS C E.
PA (CASH/) CASMAN S J.
PA (CHAP/) CHAPOVAL A.
PA (EDIN/) EDINGER S R.
PA (GERL/) GERLACH V.
PA (GORM/) GORMAN L.
PA (GUNT/) GUNTHER E.
PA (GUOX/) GUO X S.
PA (KEKU/) KEKUDA R.
PA (LEPL/) LEPLEY D M.
PA (LILL/) LI L.
PA (LIUX/) LIU X.
PA (MALV/) MALYANKAR U M.
PA (MILL/) MILLER C E.
PA (MILL/) MILLET I.
PA (PADI/) PADIGARU M.
PA (PATT/) PATTURAJAN M.
PA (PENA/) PENA C E A.
PA (RIEG/) RIEGER D K.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHIMKETS R A.
PA (SPYT/) SPYTEK K A.
PA (TAUP/) TAUPIER R J.
PA (VERN/) VERNET C A M.
PA (VOSS/) VOSS E Z.
PA (ZERH/) ZERHUSEN B D.
XX
PI Alsobrook JF, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
PI Chapoval A, Edinger SR, Gerlach V, Gorman L, Gunther E, Guo XS;
PI Kekuda R, Lepley DM, Li L, Liu X, Malyankar UM, Miller CE;
PI Millet I, Padigar M, Patturajan M, Pena CEA, Rieger DK, Shenoy SG;
PI Shimkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ;
PI Zerhusen BD;
XX WPI; 2004-122037/12.
DR P-PSDB; ADO08343.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes,
PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
PT or pharmacogenomics.
XX Claim 20; SEQ ID NO 89; 219pp; English.
XX
CC The invention relates to human NOVX polypeptides and the polynucleotides
CC encoding them. The polypeptides, polynucleotides and antibodies that bind
CC immunospecifically to the polypeptides are useful in diagnosing, treating
CC or preventing NOVX-associated disorders such as cardiomyopathy,
CC atherosclerosis, hypertension, scleroderma, obesity, cancer, diabetes,
CC haemophilia, graft-versus-host disease, AIDS, asthma, Crohn's disease,
CC multiple sclerosis, infections, anorexia, cancer-associated cachexia,
CC neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's
CC disease), haematopoietic disorders and wasting disorders. The
CC polynucleotides are also used as hybridisation probes, in chromosome
CC mapping and in tissue typing. The polypeptides are also useful as
CC vaccines. This sequence represents a human NOVX polynucleotide of the
CC invention.
XX
SQ Sequence 944 BP; 209 A; 265 C; 286 G; 184 T; 0 U; 0 Other;
Query Match 49.2%; Score 717.6; DB 12; Length 944;
Best Local Similarity 87.2%; Pred. No. 1.8e-175;
Matches 933; Conservative 0; Mismatches 4; Indels 133; Gaps 8;
QY 213 CTCCTATTTCAGCTGTGACACCTCAGAGCGCTGTGCCCAAGCATGACAGACGATG 272
Db 1 CTCCTATTTCAGCTGTGACACCTCAGAGCGCTGTGCCCTAAGCATGACAGACGATG 60
QY 273 AAAAATTCCAGTACTTGGAGAAATAAGGTGAAAGTCCAGGGGTTTAAAAATGGGCCTTC 332
Db 61 AAAAATTCCAGTACTTGGAGAAATAAGGTGAAAGTCCAGGGGTTTAAAAATGGGCCTTC 120
QY 333 CTCTCAGTCCCTCTCTGCAGCGTCTCGGCTCTGGGCGCTGCCATCTCTCTGCTCCCTGG 392

```

Db 121 |||||CTCTCAGTCCCTCTCGACGGCTCTGCTCTGGGCCCTGCCATCTCTCGCTGT-----174

Qy 393 GCCTCGGCTGCTGCTGCTGCTCATCATCTGTGTGTGTTGATTCCAAATTC 452

Db 175 -----174

Qy 453 AGAGGAGCCTGGTGACCTCGAGAACAGATTTTAGCAACTTCACCTCAAAACACTGTGGCGG 512

Db 175 -----ACTGTGGCG 184

Qy 513 AGATCAGGCACTGACTTCCAGGGCAGCAGCTTGGAGAAACGATAGCATCTCTGAAG 572

Db 185 AGATCAGGCACTGACTTCCAGGGCAGCAGCTTGGAGAAACGATAGCATCTCTGAAG 244

Qy 573 CTGAGGTGAGGGTTTCAAGCAGAGAACGCGCAGCGGGTATCTGAGCTCCAGGAACACA 632

Db 245 CTGAGGTGAGGGTTTCAAGCAGAGAACGCGCAGCGGGTATCTGAGCTCCAGGAACACA 304

Qy 633 CTACGAGAGGCACACCTAGGCCACTGTCCCACCTGCCCATCTGTGTGTGTCCTCCAGTTTC 692

Db 305 CTACGAGAGGCACACCTAGGCCACTGTCCCACCTGCCCATCTGTGTGTGTCCTCCAGTTTC 364

Qy 693 ATTCTGAATGCTCTCGAGTCCAGAGTCCAGCAGTGTGCAAGCACTTGAAGAACTGACTGCTCC 752

Db 365 ATTCTGAATGCTCTCGAGTCCAGAGTCCAGCAGTGTGTC--GACCTGAAGAACTGACTGCTCC 422

Qy 753 AGTGGCTACTCTCAACAAT-----GCTCCACTGAAGGACCTGCTGCCCGTC 805

Db 423 AGTGGCTACTCTCAACAATAGTGGGGCTCCACTGAAGGACCTGCTGCCCTGTC 482

Qy 806 AACTGGGTGGAGCACCAAGACAGTCTACTTGTCTCTCACTCTGGGATGCTCTGGGCC 865

Db 483 AACTGGGTGGAGCACCAAGACAGTCTACTTGTCTCTCACTCTGGGATGCTCTGGGCC 542

Qy 866 GAGGCTGAGAAGTACTGCGAGCTGAAGACGCCCACTCGTGGTGTATCAACTCCAGGGAG 925

Db 543 GAGGCTGAGAAGTACTGCGAGCTG-----GCGGCCACTCGTGGTGTATC-ACTCCAGGGAG 597

Qy 926 GAGCAATTTTGTCCAGAAATATCTAGGCTCGCATACCTTGGATGGCCCTCAGTGAC 985

Db 598 GAGCAATTTTGTCCAG-AAATATCTAGGCTCGCATACCTTGGATGGCCCTCAGTGAC 656

Qy 986 CTTGAAGGAGCCTGGAGTGGGTGGATGGAACAGACTATCGACCGGCTTCCAGAACTGG 1045

Db 657 CTTGAAGGAGCCTGGAGTGGGTGGATGGAACAGACTATCGACCGGCTTCCAGAACTGG 716

Qy 1046 AGCCAGGCGCAGCAGACACTGGCAGGGGCACGGGCTGGGTGGAGGCGAGACTGTGCT 1105

Db 717 AAGCCAGGCGCAGCAGACACTGGCAGGGGCACGGGCTGGGTGGAGGCGAGACTGTGCT 776

Qy 1106 CACTTCCATCCAGACGCGAGGTGGAATGACGAGCTGTGCGAGGGCCCTACCACTGGGTC 1165

Db 777 CACTTCCATCCAGACGCGAGGTGG-ATGACGAGCTGTGCGAGGGCCCTACCACTGGGTC 835

Qy 1166 TCGAGGCTGGCTGGGTGAGACCGCAGGAGTCACTGAGCTGCCCTTTGGTGGGACC 1225

Db 836 TCGAGGCTGGCTGGGTGAGACCGCAGGAGTCACTGAGCTGCCCTTTGGTGGGACC 895

Qy 1226 ACCCGCCACAGAAATGGCGGTGGGAGGAGTCTTCTCAGACCTCTT 1275

Db 896 ACCCGCCACAG-AAATGGCGGTGGGAGGAGTCTTCTCAGACCTCT 944

RESULT 7

AC10154

ID AC10154 standard; cDNA; 820 BP.

XX AC10154;

XX AC10154;

DT 02-JUN-2003 (first entry)

XX Human NOVX polynucleotide #44.

XX Human; NOVX; gene; ss; metabolic disorder; diabetes; infectious disease; obesity; anorexia; cancer; cardiovascular disorder; asthma; neurogenesis; neurodegenerative disorder; epilepsy; immune disorder; osteoarthritis; haematopoietic disorder; inflammatory skin disorder; dyslipidemia; haematopoiesis; wound healing; angiogenesis; bacterial infection; viral infection; fungal infection; helminthic infection; atherosclerosis; protozoal infection; hypertension.

OS Homo sapiens.

XX WO200290504-A2.

XX 14-NOV-2002.

PD 02-MAY-2002; 2002WO-US014342.

PF 03-MAY-2001; 2001US-0288395P.

PR 04-MAY-2001; 2001US-0288900P.

PR 07-MAY-2001; 2001US-0289087P.

PR 14-MAY-2001; 2001US-0290753P.

PR 15-MAY-2001; 2001US-0291189P.

PR 16-MAY-2001; 2001US-0291243P.

PR 18-MAY-2001; 2001US-0292001P.

PR 21-MAY-2001; 2001US-0292374P.

PR 22-MAY-2001; 2001US-0292587P.

PR 23-MAY-2001; 2001US-0293107P.

PR 29-MAY-2001; 2001US-0294110P.

PR 30-MAY-2001; 2001US-0294434P.

PR 31-MAY-2001; 2001US-0294827P.

PR 18-JUN-2001; 2001US-0298988P.

PR 31-JUL-2001; 2001US-0308901P.

PR 17-AUG-2001; 2001US-0313388P.

PR 21-AUG-2001; 2001US-0313851P.

PR 21-AUG-2001; 2001US-0313937P.

PR 17-SEP-2001; 2001US-0322701P.

PR 17-SEP-2001; 2001US-0322802P.

PR 25-SEP-2001; 2001US-0324757P.

PR 27-SEP-2001; 2001US-0325314P.

PR 27-SEP-2001; 2001US-0325682P.

PR 21-NOV-2001; 2001US-0332129P.

PR 03-DEC-2001; 2001US-0336682P.

PR 14-DEC-2001; 2001US-0340305P.

PR 01-MAY-2002; 2002US-00138588.

PA (CURA-) CURAGEN CORP.

XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ; Chapoval A, Edinger S, Gerlach V, Gorman L, Gunther E, Guo X; Kekuda R, Lepley DM, Li L, Liu X, Malyankar UM, Miller CE; Millet I, Padigaru M, Patturajan M, Pena CE, Rieger DK, Shenoy SG; Shinkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ; Zerhusen BD;

XX WPI; 2003-103512/09.

DR P-PSDB; ABU69169.

XX New isolated NOVX polypeptides and polynucleotides, useful for preventing, diagnosing or treating NOVX-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, asthma, or infections.

PS Claim 20; Page 169; 340pp; English.

XX The invention relates to human NOVX polypeptides and the polynucleotides encoding them. The polypeptides, polynucleotides and antibodies that bind immunospecifically to the polypeptides are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder. The sequences are useful for treating, preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular disorders (e.g. hypertension, atherosclerosis), neurodegenerative disorders (e.g.

PA (LEPL/) LEPLEY D M.
 PA (LILL/) LI L.
 PA (LIUX/) LIU X.
 PA (MALV/) MALVANKAR U M.
 PA (MILL/) MILLER C E.
 PA (MILL/) MILLET I.
 PA (PADI/) PADIGARU M.
 PA (PATT/) PATTURAJAN M.
 PA (PENA/) PENA C E A.
 PA (RIEG/) RIEGER D K.
 PA (SHEN/) SHENOY S G.
 PA (SHIM/) SHIMKETS R A.
 PA (SPYT/) SPYTEK K A.
 PA (TAUP/) TAUPIER R J.
 PA (VERN/) VERNET C A M.
 PA (VOSS/) VOSS E Z.
 PA (ZERH/) ZERHUSEN B D.
 XX
 PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
 PI Chapoval A, Edinger SR, Gerlach V, Gorman L, Gunther E, Guo XS;
 PI Kekuda R, Lepley DM, Li L, Liu X, Malvankar UM, Miller CE;
 PI Millet I, Padigar M, Patturajan M, Pena CE, Rieger DK, Shenoy SG;
 PI Shimkete RA, Spytek KA, Taupier RJ, Vernet CM, Voss EZ;
 PI Zerhusen BD;
 XX
 DR WPI; 2004-122037/12.
 DR P-PSDB; ADO08341.
 XX
 PT New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, diabetes,
 PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
 PT or pharmacogenomics.
 XX
 PS Claim 20; SEQ ID NO 87; 219pp; English.
 XX
 CC The invention relates to human NOVX polypeptides and the polynucleotides
 CC encoding them. The polypeptides, polynucleotides and antibodies that bind
 CC immunospecifically to the polypeptides are useful in diagnosing, treating
 CC or preventing NOVX-associated disorders such as cardiomyopathy,
 CC atherosclerosis, hypertension, scleroderma, obesity, cancer, diabetes,
 CC haemophilia, graft-versus-host disease, AIDS, asthma, Crohn's disease,
 CC multiple sclerosis, infections, anorexia, cancer-associated cachexia,
 CC neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's
 CC disease), haematopoietic disorders and wasting disorders. The
 CC polynucleotides are also used as hybridisation probes, in chromosome
 CC mapping and in tissue typing. The polypeptides are also useful as
 CC vaccines. This sequence represents a human NOVX polynucleotide of the
 CC invention.
 XX
 SQ Sequence 820 BP; 181 A; 227 C; 243 G; 169 T; 0 U; 0 Other;
 Query Match 49.0%; Score 714.6; DB 12; Length 820;
 Best Local Similarity 90.0%; Pred. No. 1e-174;
 Matches 811; Conservative 0; Mismatches 9; Indels 81; Gaps 1;
 321 ATGGGCGACATCTCTCCAGTCCCTCTCTGAGGGTCTCGGCTCTGGGCGCTGCAATCTCC 380
 1 ATGGGCGACATCTCTCTCCAGTCCCTCTCTGAGGGTCTCTGCTCTGGGCGCTGCAATCTCC 60
 381 TGTGTGCTTCCCTGGGCGCTCGGCTCTCTCTGAGGGTCTCGGCTCTGGGCGCTGCAATCTCC 440
 61 TGTGTGCTTCCCTGGGCGCTCGGCTCTCTCTGAGGGTCTCTGCTCTGGGCGCTGCAATCTCC 120
 441 ATTCCAAATTTACAGGGACCTGGTGCCTGAGACAGATTTTACCACTTCACTCA 500
 121 ATTCCAAATTTACAGGGACCTGGTGCCTGAGACAGATTTTACCACTTCACTCA 180
 501 ACACTGTGGCGAGATCCAGGCACTGACTTCCAGGGCGAGCAGCTTGGAGAAACGATAG 560
 181 ACACTGTGGCGAGATCCAGGCACTGACTTCCAGGGCGAGCAGCTTGGAGAAACGATAG 240
 561 CATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAACGGCAGGCGGTATCTGAGC 620
 |||||||

Db 241 CATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAACGCGAGG----- 285
 Qy 621 TCCAGGAACACACTACGCGAGAAGGCACACCTAGGCCACTGTCCCACACTCTGTGTGT 680
 Db 286 ----- 285
 Qy 681 GTGTCCAGTTTCATTCTGAAATGCTCTCGAGTCCAGCAGCTGTGTGCAAGACCTGAAGA 740
 Db 286 -----CAGTTTCATTCTGAAATGCTCTCGAGTCCAGCAGCTGTGTGCAAGACCTGAAGA 339
 Qy 741 AACTGACCTGCCAGGTGGCTCTCTCAACAACATGCTCCTCAAGGGGACCTGCTGCC 800
 Db 340 AACTGACCTGCCAGGTGGCTCTCTCAACAACATGCTCCTCAAGGGGACCTGCTGCC 399
 Qy 801 CGTCAACTGGGTGGAGCACCAGACAGCTGCTACTCTGTTCTCTCACTCTCGGATGTCT 860
 Db 400 CTGTCAACTGGGTGGAGCACCAGACAGCTGCTACTCTGTTCTCTCACTCTGGGATGTCT 459
 Qy 861 GGGCCGAGGCTGAGAAGTACTGCTCCAGCTGAAGAACGCCACCTGGTGTGTCATCAACTCCA 920
 Db 460 GGGCCGAGGCTGAGAAGTACTGCTCCAGCTGAAGAACGCCACCTGGTGTGTCATCAACTCCA 519
 Qy 921 GGGAGGACAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCA 980
 Db 520 GGGAGGACAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCA 579
 Qy 981 GTGACCTGAAAGGAGCTGGAGTGGGTGGATGGAACAGACTATGCGACCGGCTTCCAGA 1040
 Db 580 GTGACCTGAAAGGAGCTGGAGTGGGTGGATGGAACAGACTATGCGACCGGCTTCCAGA 639
 Qy 1041 ACTGGAAGCCAGGCGCAGCAGACGACTGCGAGGGGCGACGGGCTGGGTGGAGCGAGGACT 1100
 Db 640 ACTGGAAGCCAGGCGCAGCAGACGACTGCGAGGGGCGACGGGCTGGGTGGAGCGAGGACT 699
 Qy 1101 GTGCTCACTTCATCCAGCGGAGTGGATGAGACAGCTGTGCCAGAGGCGCTTACCCT 1160
 Db 700 GTGCTCACTTCATCCAGCTGGAGTGGATGAGACAGCTGTGCCAGAGGCGCTTACCCT 759
 Qy 1161 GGGTCTGCGAGGCTGGCTGGGTGAGACAGCAGCAGGAGAGTCACTGAGCTGCTTTGGTG 1220
 Db 760 GGGTCTGCGAGGCTGGCTGGGTGAGACAGCAGCAGGAGAGTCACTGAGCTGCTTTGGTG 819
 Qy 1221 G 1221
 Db 820 G 820
 RESULT 9
 AAF44889
 ID AAF44889 standard; cDNA; 661 BP.
 XX
 AC AAF44889;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE Human breast cancer related protein coding sequence SEQ ID NO: 45.
 XX
 KW Human; breast cancer; diagnosis; therapy; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200078960-A2.
 XX
 XX 28-DEC-2000.
 PD
 XX
 PF 23-JUN-2000; 2000WO-US017536.
 XX
 PR 23-JUN-1999; 99US-0140903P.
 PR 12-OCT-1999; 99US-0158980P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Yuqiu J, Mitcham JL;

XX WPI; 2001-041426/05.
XX New polynucleotides encoding breast tumor specific proteins, useful for
PT prevention, treatment and diagnosis of breast cancer.
XX
XX Claim 25; Page 133-134; 165pp; English.
XX
CC The present invention provides the coding sequences for a number of
CC breast cancer related proteins. These can be used in vaccinations
CC against, diagnosis of and treatment of cancer, particularly breast cancer
XX
SQ Sequence 661 BP; 156 A; 199 C; 166 G; 139 T; 0 U; 1 Other;

Query Match 41.6%; Score 606.6; DB 4; Length 661;
Best Local Similarity 98.2%; Pred. No. 9.6e-147;
Matches 612; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
QY 1 GTTGAGGAGATGGATGCTCCAGATGATAGGCTCTCTGGGATTTTCAGACCCAGACGAGC 60
Db |||||||
QY 61 AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGGCTCCCAACTCTGAGTGAC 120
Db |||||||
QY 121 GTCCCACTCTGCTCTTGGAGCACAACCAACGTTGGGAATCACACCTCCAGACCTCCCA 180
Db |||||||
QY 181 CAGCTCCACCCAGATGCGGGCGGCGCTCCCTCCATTTTCAGTGTGCAACTCAGAG 240
Db |||||||
QY 241 CCCTGTTGGCCCAAGCATGACAGGACGATGATAAACTTCCAGTACTTGGAGATAAGGT 300
Db |||||||
QY 301 GAAAGTCAGGGGTTTAAATAATGGCCACTTCTCTCCAGTCTCTCTGAGCGTCTCCG 360
Db |||||||
QY 361 CTCTGGCCCTGCGATCTCTGCTGCTCCCTGGGCTGCTGCTGCTGCTGCTGCTCAT 420
Db |||||||
QY 421 CTGTGTGTTGGATTCCAAATTTCCAAATTTCCAGAGGACCTGTGTGACCTTGAGACAGA 480
Db |||||||
QY 481 TTTTAGCACTTCACTTCAACACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGGAG 540
Db |||||||
QY 541 CAGCTTGGAGAAACGATAGCATCTCTGAAGCTGAGTGGAGGTTTCAAGCAGGAGACG 600
Db |||||||
QY 601 GCAGGCGAGGGGTATCTCAGCTCC 623
Db |||||||
QY 601 GCAGGCGAGGGGTATCTCAGAAATGC 623

RESULT 10
ID ABT42075 standard; DNA; 1358 BP.
XX
AC ABT42075;
XX
DT 26-JUN-2003 (first entry)
XX
DE Toxicity modelling related rat gene SEQ ID No 1777.
XX

KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KW database; drug screening; toxicity assay; rat; db.
XX
OS Rattus norvegicus.
XX
PN WO200295000-A2.
XX
PD 28-NOV-2002.
XX
XX 22-MAY-2002; 2002WO-US016173.
XX
XX 22-MAY-2001; 2001US-0292335P.
XX 13-JUN-2001; 2001US-0297523P.
XX 13-JUN-2001; 2001US-0298252P.
XX 10-JUL-2001; 2001US-0303807P.
XX 10-JUL-2001; 2001US-0303808P.
XX 10-JUL-2001; 2001US-0303810P.
XX 28-AUG-2001; 2001US-0315047P.
XX 27-SEP-2001; 2001US-0324928P.
XX 22-OCT-2001; 2001US-0330462P.
XX 01-NOV-2001; 2001US-0330867P.
XX 21-NOV-2001; 2001US-0331805P.
XX 06-DEC-2001; 2001US-0336144P.
XX 19-DEC-2001; 2001US-0340873P.
XX 21-FEB-2002; 2002US-0357842P.
XX 21-FEB-2002; 2002US-0357843P.
XX 21-FEB-2002; 2002US-0357844P.
XX 15-MAR-2002; 2002US-0364134P.
XX 08-APR-2002; 2002US-0370144P.
XX 08-APR-2002; 2002US-0370206P.
XX 17-APR-2002; 2002US-0372794P.
XX 21-APR-2002; 2002US-0371679P.
XX (GENE-) GENE LOGIC INC.
XX Mendrick D, Porter M, Johnson K, Higgs B, Caastle A, Elashoff M;
XX WPI; 2003-148464/14.
XX
XX Predicting at least one toxic effect of a compound, useful for toxicity
PT modeling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
XX Example 4; Page; 446pp; English.
XX
XX The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX
SQ Sequence 1358 BP; 340 A; 325 C; 365 G; 328 T; 0 U; 0 Other;

Query Match 30.0%; Score 438; DB 10; Length 1358;
Best Local Similarity 68.1%; Pred. No. 7.4e-103;
Matches 676; Conservative 0; Mismatches 295; Indels 21; Gaps 4;
QY 210 TGCTCCATTTTCAGTGTGACCAACCTCAGAGCC--GTGTTGGCCCAAGCATGACAAGGA 266
Db |||||||
Db 122 TGCTCGTTTCAGTGTGACCAACCTCAGAGCCCTTAGAGGCTGTGTTGTGCAAGCATGACAATGG 181

Db 581 TGGGAAGAAGATGAAGTCGCTAGAGTCCAGCTGGAGAAACAGCAGAGAGGACCTGAGTG 640
Qy 606 CAGGGGTATCTGAGCTCCAGGAACACACTACGAGAGGACACACCTAGGCCACTGCCCC 665
Db 641 AAGATCACTCCAGCTGCTGCTCCAGCTGAAGCAGTTCGTGTGACCTGCGGAGCCTGA 700
Qy 666 ACTGCCCATCTGTGTGTGCCAGTTTCATTCTGAATGCTCTCGAGTCCAGCAGCTGG 725
Db 701 GCTGTGATGCGGCGCTCCAGGGCAATGGTAAGGAGGACCGCGGCCGCTCTCTGC 760
Qy 726 TGAAGACCTGAAGAACTGACTGCTGCCAGGTGCTACTCTCAACAACTGCTTCCACTG 785
Db 761 CTCGCCCTTCTCTGGGCGAGCTTACGCCCCCTGCGGCCGCTTCTCCGCTCAGGCTCAG 820
Qy 786 AAGGAGCTGCTGCCCCGCTCACTGGGTGGAGCACCAGACAGCTGCTACTGGTCTCTC 845
Db 821 AAGGACCTGCTGCCCCGCTCACTGGGTGGAGCAGCGGAGCTGCTACTGGTCTCTC 880
Qy 846 ACTCTGGGATGCTCTGGGCGAGGCTTGAGAACTACTGCGAGCTGAAGAACGCCACCTGG 905
Db 881 GCTCCGGGAAGCTTGGGTGACCGCGACACTACTGCGGCTGGAGGACGCGACCTGG 940
Qy 906 TGGTCATCACTCCAGGAGGAGCAGAAATTTGTTCAGAAATATCTAGGCTCCGCATACA 965
Db 941 TGGTGTCACTGCTCGGAGGAGCAGAAATTTGTCCAGCACCACATAGGCCCTGTGAACA 1000
Qy 966 CTTGGATGGGCTCAGTGACCTTGAGGAGCCTGGAAGCTGGAGTGGATGAACAGACTATG 1025
Db 1001 CTTGGATGGGCTTCCAGCACCAGGCGCCCTTGAAGTGGGTGGACGGGACTAG 1060
Qy 1026 CGACCGGCTTCCAGAACTGGAGCCAGGCGCAGCAGCAGCTGGCAGGGGCGCGGCTGG 1085
Db 1061 AGACGGGCTTCAAGAACTGGAGCCGAGCAGCGCAGCAGCTGGTACGCGCAGGGCTG 1120
Qy 1086 GTGGAGCGAGGACTGTGTCTCACTTCCATCCAGACGCGAGGTGGAATGACGCTTGCC 1145
Db 1121 GAGGAGCGAGGACTGTGCCCACTTCAACGACGCGCGCTTGAAGCAGCAGCAGCTTGCC 1180
Qy 1146 AGAGGCGCTTACCACTGGGTCTGGAGGCTGGGCTGGGTGAGACCGACCGAGGAG 1198
Db 1181 AGAGGCGCTTACCGCTGGGTCTGGAGCAGAGCTGGACAAAGGCGCAGCGAGGAG 1233

RESULT 13

ABN95732
ID ABN95732 standard; DNA; 1277 BP.
XX
AC ABN95732;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2230 used to diagnose liver cancer.
XX
KW Gene; liver cancer; db; hepatocellular carcinoma; hepatotropic;
KW Metastatic liver tumor; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-05030589.
XX
PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
DR

XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX
PS Claim 1; SEQ ID NO 2230; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 1277 BP; 271 A; 394 C; 378 G; 234 T; 0 U; 0 Other;

Query Match 26.2%; Score 382; DB 6; Length 1277;
Best Local Similarity 66.1%; Pred. No. 2.4e-88;
Matches 546; Conservative 0; Mismatches 245; Indels 87; Gaps 3;

Qy 224 CTGTGACAACTCAGAGCCGTGTTGGCCCAAGCAGTGAACAAGAGCGTATGAACATTCACG 283
Db 140 CTGAGCAATCCAGGTCAGCGCCAGCCCTATCATGACCAAGAGGATATCAAGACCTTCAG 199
Qy 284 TACTTGGAGATAGGTGAAGTCCAG---GGGTTTAAATATGGCCACTTCTCTCCACG 340
Db 200 CATCTGGCAATAGGAGAGTGAACCAATCAGCTCAGAAAGGCGCACCTCTCTCCACG 259
Qy 341 TCCCTCTCGACGCTCTCCGCTCTGGGCGCTGCGCATCTCTGCTGTCTCCCTGGGCTCCGC 400
Db 260 CCCCTCTCGACGCTCTCTGCTCGGACCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 319
Qy 401 CTGCTGCTGCTGCTCATCATCTGTGTGTGGTTCGAAATTCGAAATTCAGAGGAC 460
Db 320 CTCCTGCTGCTGT 379
Qy 461 CTGTGTGACCTCGAGAACAGATTTTAGCAACTTCACTCAAACTGTGCGGAGATCCAG 520
Db 380 CTGGGGGCTCGAGAGAGACGTTTACGCAACTTTCAGCGAGCAGGAGCCCGAGGTCAAG 439
Qy 521 GCACCTGACTTCCAGGGCAGCAGCTTGGAGAAACGATAGCATCTCTGAAAGCTCAGGTG 580
Db 440 GGCTTGAGCACCAGGAGGCAATCTGGGAGAGAGATGAGTGCCTAGAGTCCAGCTG 499
Qy 581 GAGGTTTTCAAGCAGGAACCGCAGCGGGGTATCTGAGTCCAGGAAACACACTACCGAG 640
Db 500 GAGAAACAGCAGAGGACCTGAGTG----- 524
Qy 641 AAGGCACACTAGGCCACTGTCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 700
Db 525 -----AAGATCACTCCAGC 538
Qy 701 ATGCTCTCGAGTCCAGCAGCTGGTGAAGACCTGGAAGAACTGACCTGCGCAGGTGGCT 760
Db 539 CTGCTGCTCCAGCTGAAGCAGTTCGTGTCTGACCTGCGGAGCCTGAGCTGTCTGAGTGGG 598
Qy 761 ACTCTCAACAACTGCTTCCACTGAAGGAGCCTGCTGCCCGCTCAACTGGGTGAGAGC 820
Db 599 GCGCTCCAGGGCAATGGCT---CAGAAAGGAGCCTGCTGCCCGGTCAACTGGGTGAGAGC 655
Qy 821 CAAGACAGCTGCTACTGGTCTCTCTCACTCTGGAGTGTCTCTGGGCGGAGGCTGAGAGTAC 880
Db 656 GAGCGCAGCTGCTACTGGTCTCTCTCGCTCCGGAGAGGCTGGGCTGACGCCGCAACTAC 715

Db 956 GGCCGCTGGAAACGACGACGCTCTGCCAGAGGCCCTACCGCTGGGTCTGCCGAGACAGACTG 1015
Qy 1181 GGTCAGACCCAGCCAGGAG 1198
Db 1016 GACAGGCCCCAGCCAGGAG 1033

Search completed: May 27, 2005, 20:40:05
Job time : 848 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 22:27:31 ; Search time 922 Seconds
(without alignments)
9708.761 Million cell updates/sec

Title: US-10-829-107-3
Perfect score: 1458
Sequence: 1 gttgaggagatgggatgtcc.....taataaaatgcactcattt 1458

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5700845 seqs, 3069779757 residues

Total number of hits satisfying chosen parameters: 11401690

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1458	100.0	1458	9	US-09-862-802-3	Sequence 3, Appli
1458	100.0	1458	18	US-10-829-107-3	Sequence 3, Appli
1434.2	98.4	1617	9	US-09-728-952-45	Sequence 45, Appl
1218	83.5	1344	10	US-09-960-706-492	Sequence 492, App
1108	76.0	1370	9	US-09-862-802-9	Sequence 9, Appli
1108	76.0	1370	18	US-10-829-107-9	Sequence 9, Appli
1039.2	71.3	1072	17	US-10-138-588-89	Sequence 89, Appl
724.2	49.7	929	9	US-09-925-302-18	Sequence 18, Appl
724.2	49.7	929	10	US-09-925-302-18	Sequence 18, Appl
714.6	49.0	820	17	US-10-138-588-87	Sequence 87, Appl
438	39.0	1358	17	US-10-152-319A-1777	Sequence 1777, Ap

12 382 26.2 1277 9 US-09-880-107-2230 Sequence 2230, Ap
13 382 26.2 1277 17 US-10-305-720-1186 Sequence 1186, Ap
14 321.4 22.0 1055 17 US-10-236-392-9 Sequence 9, Appli
15 321.4 22.0 1300 9 US-09-880-107-3731 Sequence 3731, Ap
16 303 20.8 1112 17 US-10-236-392-7 Sequence 7, Appli
17 282.8 19.4 1309 9 US-09-880-107-2235 Sequence 2235, Ap
18 282.8 19.4 1309 17 US-10-236-392-5 Sequence 5, Appli
19 282.8 19.4 1309 18 US-10-283-975A-169 Sequence 169, App
20 246 16.9 443 9 US-09-960-352-7063 Sequence 7063, Ap
21 231.8 15.3 466 10 US-09-918-995-8784 Sequence 8784, Ap
22 223.6 15.3 1430 17 US-10-062-674-1931 Sequence 1931, Ap
23 212.4 14.6 1290 9 US-09-917-800A-1642 Sequence 1642, Ap
24 183 12.6 400 17 US-10-170-097-46 Sequence 46, Appl
25 183 12.6 400 19 US-10-926-684-46 Sequence 46, Appl
26 155.6 10.7 448 17 US-10-170-097-49 Sequence 49, Appl
27 155.6 10.7 448 19 US-10-926-684-49 Sequence 49, Appl
28 154.4 10.6 448 17 US-10-170-097-51 Sequence 51, Appl
29 154.4 10.6 448 19 US-10-170-097-571 Sequence 571, Appl
30 154.4 10.6 448 17 US-10-926-684-51 Sequence 51, Appl
31 154.4 10.6 448 19 US-10-926-684-571 Sequence 571, Appl
32 154 10.6 448 17 US-10-170-097-50 Sequence 50, Appl
33 154 10.6 448 19 US-10-926-684-50 Sequence 50, Appl
34 146 10.0 402 9 US-09-960-352-9823 Sequence 9823, Ap
35 142.4 9.8 410 9 US-09-960-352-11324 Sequence 11324, A
36 142 9.7 463 10 US-09-918-995-785 Sequence 785, App
37 126.2 8.7 220 17 US-10-062-674-1353 Sequence 1353, Ap
38 118 8.1 761 13 US-10-027-632-17618 Sequence 17618, A
39 118 8.1 761 13 US-10-027-632-151016 Sequence 151016, A
40 118 8.1 761 17 US-10-027-632-17618 Sequence 17618, A
41 118 8.1 761 17 US-10-027-632-151016 Sequence 151016, A
42 116.8 8.0 421 10 US-09-918-995-8213 Sequence 8213, Ap
43 115.4 7.9 402 9 US-09-960-352-2730 Sequence 2730, Ap
44 114.2 7.8 482 10 US-09-918-995-30378 Sequence 30378, A
45 109 7.5 401 9 US-09-960-352-6627 Sequence 6627, Ap

ALIGNMENTS

RESULT 1
US-09-862-802-3
; Sequence 3, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid
; NAME/KEY: CDS
; LOCATION: 257...1204
; OTHER INFORMATION: protein coding sequence
; NAME/KEY: variation
; LOCATION: 608...673
; OTHER INFORMATION: short form lacks these nucleotides
US-09-862-802-3

Query Match 100.0%; Score 1458; DB 9; Length 1458;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTGAGAGATGGGATGCCAGATAGGCTCTGGGATTCAGACCCAGACG 60
|||||

Db 1 GTTAGGAGATGGATGTCCCAAGATAGAGGCTCTGGGATTTCCAGACCCCAAGACCGC 60
QY 61 AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTGAC 120
Db 61 AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTGAC 120
QY 121 GTCCCACTCTGTCTCTGAGACACAAACAGTGGGAATCAACCTCCAGACCTCCCA 180
Db 121 GTCCCACTCTGTCTCTGAGACACAAACAGTGGGAATCAACCTCCAGACCTCCCA 180
QY 181 CAGCTCCACCCAGACTGGCGCGCGCTGCTCCATTTTCAGCTGTGACCACTCCAG 240
Db 181 CAGCTCCACCCAGACTGGCGCGCGCTGCTCCATTTTCAGCTGTGACCACTCCAG 240
QY 241 CCGTGTGGCCCAAGCATCAAGGACGTATGAAAATTTCCAGTACTTGGAGAAATAGGT 300
Db 241 CCGTGTGGCCCAAGCATCAAGGACGTATGAAAATTTCCAGTACTTGGAGAAATAGGT 300
QY 301 GAAAGTCCAGGGGTTTAAAATGGGCCACTTCCCTCTCCAGTCCCTCTCGACGGTCTCG 360
Db 301 GAAAGTCCAGGGGTTTAAAATGGGCCACTTCCCTCTCCAGTCCCTCTCGACGGTCTCG 360
QY 361 CTCCTGGCCCTGCCATCTCTGTGTCTCTGGGCTGGGCTGCTGCTGCTGCTCAT 420
Db 361 CTCCTGGCCCTGCCATCTCTGTGTCTCTGGGCTGGGCTGCTGCTGCTGCTCAT 420
QY 421 CTGTGTGGTTGGATTCAAAATTTCCAAAATTTCCAGAGGACCTGGTGACCTGAGAACAGA 480
Db 421 CTGTGTGGTTGGATTCAAAATTTCCAAAATTTCCAGAGGACCTGGTGACCTGAGAACAGA 480
QY 481 TTTTAGCAACTTCACCTCAAACTGTGGCGAGATCCAGGCACTGACTTCCAGGGCGAG 540
Db 481 TTTTAGCAACTTCACCTCAAACTGTGGCGAGATCCAGGCACTGACTTCCAGGGCGAG 540
QY 541 CAGCTTGGAGAAACATGACATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACG 600
Db 541 CAGCTTGGAGAAACATGACATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAACG 600
QY 601 GCAGGCAGGGGTATCTGAGCTCCAGAACACACTACGCAAGAGGCACACTAGGCCACTG 660
Db 601 GCAGGCAGGGGTATCTGAGCTCCAGAACACACTACGCAAGAGGCACACTAGGCCACTG 660
QY 661 TCCCACTGCCCATCTGTGTGTCTCCAGTTCATTTCTGAAATGCTCTCGAGTCCAGCA 720
Db 661 TCCCACTGCCCATCTGTGTGTCTCCAGTTCATTTCTGAAATGCTCTCGAGTCCAGCA 720
QY 721 GCTGTGCAAGACCTGAAGAACTGACCTGCCAGTGGCTACTCTCAACAAACAATGCCCTC 780
Db 721 GCTGTGCAAGACCTGAAGAACTGACCTGCCAGTGGCTACTCTCTCAACAAACAATGCCCTC 780
QY 781 CACTGAAGGGACCTGTGCCCCGTCAACTGGGTGGAGACCAAGACAGCTGCTACTGGTT 840
Db 781 CACTGAAGGGACCTGTGCCCCGTCAACTGGGTGGAGACCAAGACAGCTGCTACTGGTT 840
QY 841 CTCTCACTCTGGGATCTCTGGGCCGAGGTGAGAAAGTACTGCGAGCTGAAGAACGCCCA 900
Db 841 CTCTCACTCTGGGATCTCTGGGCCGAGGTGAGAAAGTACTGCGAGCTGAAGAACGCCCA 900
QY 901 CTTGTGTGTCATCAACTCAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGC 960
Db 901 CTTGTGTGTCATCAACTCAGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGC 960
QY 961 ATACACTGTAGTGGGCTCAGTGACCTCAAGGAGCTGGAAGTGGGTGGATGGAAACAGA 1020
Db 961 ATACACTGTAGTGGGCTCAGTGACCTCAAGGAGCTGGAAGTGGGTGGATGGAAACAGA 1020
QY 1021 CTATGCGACCGGCTTCAGAACTGGAAGCAGGCCAGCAGCACTGGCAGGGGACCGG 1080
Db 1021 CTATGCGACCGGCTTCAGAACTGGAAGCAGGCCAGCAGCACTGGCAGGGGACCGG 1080
QY 1081 GCTGGGTGGAGGCGAGGACTGTGCTCATCTTCATCCAGACGGCAGGTGGAATGACGACGT 1140
Db 1081 GCTGGGTGGAGGCGAGGACTGTGCTCATCTTCATCCAGACGGCAGGTGGAATGACGACGT 1140

RESULT 2

US-10-829-107-3
; Sequence 3, Application US/10829107
; Publication No. US20040192892A1
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SP0695 B
; CURRENT APPLICATION NUMBER: US/10/829,107
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/862,802A
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1458
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid and protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (257)..(1204)
; OTHER INFORMATION: protein coding sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (608)..(673)
; OTHER INFORMATION: short form lacks nucleotides 608-673
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (775)..(776)
; OTHER INFORMATION: ASGPRm (Table 2) has sequence insert encoding GEE between nucleot
; OTHER INFORMATION: ides 775-776
; NAME/KEY: misc feature
; LOCATION: (1064)..(1064)
; OTHER INFORMATION: nucleotide 1064 of DCMP2s may be A, which would encode Asn rather
; OTHER INFORMATION: than Asp at the residue numbered 270

US-10-829-107-3

Query Match 100.0%; Score 1458; DB 18; Length 1458;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GTGAGGAGATGGGATGCTCCAGATGATAGGGCTCTGGGATTTCCAGACCCAGACGAGC	60
DB	1	GTGAGGAGATGGGATGCTCCAGATGATAGGGCTCTGGGATTTCCAGACCCAGACGAGC	60
QY	61	AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGGGCTCCCAAATCTGAGTGAC	120
DB	61	AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGGGCTCCCAAATCTGAGTGAC	120
QY	121	GTCCCACTCTGTCTCTGAGACACAAACGAGTGGGAATCACACCTCCAGACCTCCCA	180
DB	121	GTCCCACTCTGTCTCTGAGACACAAACGAGTGGGAATCACACCTCCAGACCTCCCA	180
QY	181	CAGTCCACCCAGACCTGGGGCGGCGCTCCCTCCATTTTCACTGTGACAACTCAGAG	240
DB	181	CAGTCCACCCAGACCTGGGGCGGCGCTCCCTCCATTTTCACTGTGACAACTCAGAG	240
QY	241	CCGTGTTGGCCCAAGCATGACAGGAGCTATGAAAATCTCCAGTACTTGGAGAAATAGGT	300
DB	241	CCGTGTTGGCCCAAGCATGACAGGAGCTATGAAAATCTCCAGTACTTGGAGAAATAGGT	300
QY	301	GAAAGTCAGGGGTTTAAATAATGGGCCACTTCTCTCCAGTCCCTCCAGGCTCTCCG	360
DB	301	GAAAGTCAGGGGTTTAAATAATGGGCCACTTCTCTCCAGTCCCTCCAGGCTCTCCG	360
QY	361	CTCTGGGCGCTGCATCTCTGCTGTCTCCCTGGGCGCTCGGCTCTGCTGTGCTCATCAT	420
DB	361	CTCTGGGCGCTGCATCTCTGCTGTCTCCCTGGGCGCTCGGCTCTGCTGTGCTCATCAT	420
QY	421	CTGTGTGGTGGATTTCCAAATTTCCAGAGGAGCTGTGTGACCTCTGAGAACAGA	480
DB	421	CTGTGTGGTGGATTTCCAAATTTCCAGAGGAGCTGTGTGACCTCTGAGAACAGA	480
QY	481	TTTTAGCAATTCACCTCAAACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGCAG	540
DB	481	TTTTAGCAATTCACCTCAAACTGTGGCGGAGATCCAGGCACTGACTTCCAGGGCAG	540
QY	541	CAGTTGGAGAAACGATGATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAGC	600
DB	541	CAGTTGGAGAAACGATGATCTCTGAAAGCTGAGGTGGAGGGTTTCAAGCAGGAGC	600
QY	601	GCAGGACGGGTATCTGAGCTCCAGGAACACTACGACGAAGGACACCTAGGCCACTG	660
DB	601	GCAGGACGGGTATCTGAGCTCCAGGAACACTACGACGAAGGACACCTAGGCCACTG	660
QY	661	TCCCCACTGCCCATCTGTGTGTGTCAGTTCATTTCTGAAATGCTCTGCGAGTCCAGCA	720
DB	661	TCCCCACTGCCCATCTGTGTGTGTCAGTTCATTTCTGAAATGCTCTGCGAGTCCAGCA	720
QY	721	GCTGGTGCAGACCTGAGAACTGACCTGCCAGTGGCTACTCTCAACAACTAGCCCTC	780
DB	721	GCTGGTGCAGACCTGAGAACTGACCTGCCAGTGGCTACTCTCAACAACTAGCCCTC	780
QY	781	CAGTGAAGGACCTGTCTGCCCGCTCACTGGTGGAGCAACAGACAGCTGCTACTGGTT	840
DB	781	CAGTGAAGGACCTGTCTGCCCGCTCACTGGTGGAGCAACAGACAGCTGCTACTGGTT	840
QY	841	CTCTCACTCTGGGATGCTCTGGGCGGAGGCTGAGAACTGCTCCAGCTGAGAACGCCCA	900
DB	841	CTCTCACTCTGGGATGCTCTGGGCGGAGGCTGAGAACTGCTCCAGCTGAGAACGCCCA	900
QY	901	CTTGGTGGTATCACTCCAGGGAGGAGCAATTTTGTCCAGAAATATCTAGGCTCCGC	960
DB	901	CTTGGTGGTATCACTCCAGGGAGGAGCAATTTTGTCCAGAAATATCTAGGCTCCGC	960
QY	961	ATACACCTGTAGTGGGCTCAGTGCACCTGAGGAGGCTGGAAGTGGGTGGATGGAACAGA	1020
DB	961	ATACACCTGTAGTGGGCTCAGTGCACCTGAGGAGGCTGGAAGTGGGTGGATGGAACAGA	1020

QY	1021	CTATGCAACGGCTTCCAGAACTGGAAAGCCAGCCAGCAGCACTGGCAGGGGACCG	1080
DB	1021	CTATGCAACGGCTTCCAGAACTGGAAAGCCAGCCAGCAGCACTGGCAGGGGACCG	1080
QY	1081	GCTGGGTGGAGGGAGGAGCTGTCTCACTTCCATCCAGACGGCAGGTGGAAATGACGAGT	1140
DB	1081	GCTGGGTGGAGGGAGGAGCTGTCTCACTTCCATCCAGACGGCAGGTGGAAATGACGAGT	1140
QY	1141	CTGCCAGAGGCCCTTACCACTGGGTCTGCAGGCTGGCTGGGTTCAGACCCAGGAGAG	1200
DB	1141	CTGCCAGAGGCCCTTACCACTGGGTCTGCAGGCTGGCTGGGTTCAGACCCAGGAGAG	1200
QY	1201	TCACTGAGCTGCTTGTGGGACCAACCGGCCACAGAAATGCGCGTGGGAGGAGACTC	1260
DB	1201	TCACTGAGCTGCTTGTGGGACCAACCGGCCACAGAAATGCGCGTGGGAGGAGACTC	1260
QY	1261	TTCTCAGCACTCTCTCGCAAGCCGCTCTGGGAGAGAAATAAGCACTGGGAGATTGGAAG	1320
DB	1261	TTCTCAGCACTCTCTCGCAAGCCGCTCTGGGAGAGAAATAAGCACTGGGAGATTGGAAG	1320
QY	1321	CACCTGCTAACTTTGAAATTTTCTCTTTTAAATAAGATGGTATAGTGTCTTA	1380
DB	1321	CACCTGCTAACTTTGAAATTTTCTCTTTTAAATAAGATGGTATAGTGTCTTA	1380
QY	1381	AGCTTTTATTTTTCGAACTTTTGAAAGTCACTTCACTGAAGGTATATTTTACATA	1440
DB	1381	AGCTTTTATTTTTCGAACTTTTGAAAGTCACTTCACTGAAGGTATATTTTACATA	1440
QY	1441	ATAAAAATGCACATTTT 1458	
DB	1441	ATAAAAATGCACATTTT 1458	

RESULT 3

US-09-728-952-45
; Sequence 45, Application US/09728952
; Patent No. US20020111302A1

; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Vicki
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020111302A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09/728,952
; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 45
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (397) .. (1356)
US-09-728-952-45

Query Match 98.4%; Score 1434.2; DB 9; Length 1617;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

QY	1	GTGAGGAGATGGGATGCTCCAGATGATAGGGCTCTGGGATTTCCAGACCCAGACGAGC	60
DB	141	GTGAGGAGATGGGATGCTCCAGATGATAGGGCTCTGGGATTTCCAGACCCAGACGAGC	200

QY 61 AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTGAC 120
DB 201 AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGCTCCCACTCTGAGTGAC 260
QY 121 GTCCCACTCTGGTCTTGCAGGACAAACAAAGTGGGAATCAACCGCTCCAGCTCCCA 180
DB 261 GTCCCACTCTGGTCTTGCAGGACAAACAAAGTGGGAATCAACCGCTCCAGCTCCCA 320
QY 181 CAGCTCCACCCAGAGCTGGGCGCGCGCTGCTCCATTTTCAGCTGTGACAACTTCAGAG 240
DB 321 CAGCTCCACCCAGAGCTGGGCGCGCGCTGCTCCATTTTCAGCTGTGACAACTTCAGAG 380
QY 241 CCGTGTGGCCCAAGCATGACAGGACGATGAAGAACTTCCAGTACTTGGAGAAATAGGT 300
DB 381 CCGTGTGGCCCAAGCATGACAGGACGATGAAGAACTTCCAGTACTTGGAGAAATAGGT 440
QY 301 GAAAGTCCAGGGGTTTAAAAATGGGCACTTCTCTCCAGTCCCTCTGCGAGGCTCTCGG 360
DB 441 GAAAGTCCAGGGGTTTAAAAATGGGCACTTCTCTCCAGTCCCTCTGCGAGGCTCTCGG 500
QY 361 CTCTGGGCGCTGCGATCTCTGCTGTCTGCTGGGCGCTGCGGCTGCTGCTGCTGCTCAT 420
DB 501 CTCTGGGCGCTGCGATCTCTGCTGTCTGCTGGGCGCTGCGGCTGCTGCTGCTGCTCAT 560
QY 421 CTGTGTGGTGGATTCAAAAATTCAAAAATTCAGAGGACCTGGTGACCCCTGAGAACAGA 480
DB 561 CTGTGTGGTGGATTCAAAAATTCAAAAATTCAGAGGACCTGGTGACCCCTGAGAACAGA 620
QY 481 TTTTAGCACTTCACTTCAAACTGTGGGGGAGATCCAGGCACTGACTTCCAGGGGAG 540
DB 621 TTTTAGCACTTCACTTCAAACTGTGGGGGAGATCCAGGCACTGACTTCCAGGGGAG 680
QY 541 CAGTTTGAAGAAACCATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAAGC 600
DB 681 CAGTTTGAAGAAACCATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAAGC 740
QY 601 GCAGGAGGGGTATCTGAGCTCCAGGAACACACTACGAGAAAGGACACCTTAGGCCACTG 660
DB 741 GCAGGAGGGGTATCTGAGCTCCAGGAACACACTACGAGAAAGGACACCTTAGGCCACTG 800
QY 661 TCCCACTGCCCATCTGTGTGTCCAGTTTCTGAAATGCTCTCGGAGTCCAGCA 720
DB 801 TCCCACTGCCCATCTGTGTGTCCAGTTTCTGAAATGCTCTCGGAGTCCAGCA 860
QY 721 GCTGTGCAAGACCTCAAGAACTGACCTGCCAGGTGGCTACTCTCAACAAAT----- 775
DB 861 GCTGTGCAAGACCTCAAGAACTGACCTGCCAGGTGGCTACTCTCAACAAATGGTGA 920
QY 776 ----GCTTCCACTGAAGGACCTGCTGCCCGCTCAACTGGGTGGAGCACAAGACAGCTG 831
DB 921 GGAAGCCTCCACTGAAGGACCTGCTGCCCTGTCACTGGGTGGAGCACAAGACAGCTG 980
QY 832 CTACTGGTCTCTCACTCTGGGATGTCTGGGCGAGGCTGAGAGTACTGCCAGCTGAA 891
DB 981 CTACTGGTCTCTCACTCTGGGATGTCTGGGCGAGGCTGAGAGTACTGCCAGCTGAA 1040
QY 892 GAAAGCCCACTGGTGTCTCACTCACTCCAGGGAGGAGCAATTTTGTCCAGAAATATCT 951
DB 1041 GAAAGCCCACTGGTGTCTCACTCACTCCAGGGAGGAGCAATTTTGTCCAGAAATATCT 1100
QY 952 AGGCTCCGATATACCTTGGATGGGCTCAGTGACCCCTGAAGGAGCCTGGAAGTGGGTGA 1011
DB 1101 AGGCTCCGATATACCTTGGATGGGCTCAGTGACCCCTGAAGGAGCCTGGAAGTGGGTGA 1160
QY 1012 TGGAAACAGATATGCAACCGGCTTCAGAACTGGAAGCCAGGCGCAGACGACTGGCA 1071
DB 1161 TGGAAACAGATATGCAACCGGCTTCAGAACTGGAAGCCAGGCGCAGACGACTGGCA 1220
QY 1072 GGGGACAGGCTGGGTGGAGGGGAGGACTGTGCTCACTTCCATCCAGAGCGGAGGTGGA 1131
DB 1221 GGGGACAGGCTGGGTGGAGGGGAGGACTGTGCTCACTTCCATCCAGAGCGGAGGTGGA 1280
QY 1132 TGACGAGCTGTGCGAGAGGCGCTTACCACTGGGTCTGCGAGGCTGGGCTGGGTCAGACCAG 1191

DB 1281 TGACGAGCTCTGCCAGAGGCCCTTACCCTAGCTGGGTCTGCGAGGCTGGCTTGCAGACCAG 1340
QY 1192 CCAGGAGAGTCACTGAGCTGCTTTGGTGGGACCAACCGGCCACAGAAATGCGGTTGGGA 1251
DB 1341 CCAGGAGAGTCACTGAGCTGCTTTGGTGGGACCAACCGGCCACAGAAATGCGGTTGGGA 1400
QY 1252 GGAGGACTCTTCTCCAGGACCTCTCGAAGACCGCTCTGGGAGAGAAATAAGCACTGGGA 1311
DB 1401 GGAGGACTCTTCTCCAGGACCTCTCGAAGACCGCTCTGGGAGAGAAATAAGCACTGGGA 1460
QY 1312 GATTGGAAGCACTGCTAAATTTTGAATTTTCTTTTCTTTTAAATTTTAAAAAGATGATATA 1371
DB 1461 GATTGGAAGCACTGCTAAATTTTGAATTTTCTTTTCTTTTAAATTTTAAAAAGATGATATA 1520
QY 1372 GTGTCTTTAAGCTTTTATTTTCCAACTTTTCCAACTTTTGAAGTCAACTCATGAGGTATAAT 1431
DB 1521 GTGTCTTTAAGCTTTTATTTTCCAACTTTTCCAACTTTTGAAGTCAACTCATGAGGTATAAT 1580
QY 1432 TTTTACATAATAAAAAATGCACCTATTT 1458
DB 1581 TTTTACATAATAAAAAATGCACCTATTT 1607

RESULT 4
US-09-960-706-492
; Sequence 492, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; PRIOR APPLICATION NUMBER: 2001-09-24
; PRIOR FILING DATE: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 492
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 D50532
US-09-960-706-492

Query Match 83.5%; Score 1218; DB 10; Length 1344;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1328; Conservative 0; Mismatches 0; Indels 90; Gaps 2;
QY 50 CCAAGACCAGCAGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGTCCCAA 109
DB 1 CCAAGACCAGCAGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGTCCCAA 60
QY 110 CTCTGAGTGAGTCCCACTCTGGTCTTGGCAGCACACCAACGTTGGGAATCACACCTC 169
DB 61 CTCTGAGTGAGTCCCACTCTGGTCTTGGCAGCACACCAACGTTGGGAATCACACCTC 120
QY 170 CAGACTCTCCCACTCCACCCCACTGAGCTGGGCGCGGCGCTGCTCCATTTTCAGCTGTGA 229
DB 121 CAGACTCTCCCACTCCACCCCACTGAGCTGGGCGCGGCGCTGCTCCATTTTCAGCTGTGA 180
QY 230 CAACCTCAGAGCCGTTGTTGGCCCAAGCATGACAGGAGTATGAAAACTTCAGTACTTG 289
DB 181 CAACCTCAGAGCCGTTGTTGGCCCAAGCATGACAGGAGTATGAAAACTTCAGTACTTG 240
QY 290 GAGATAGGTCAAGTCCAGGGGTTTAAAAATGGGCACCTTCTCTCAGTCCCTCTG 349
DB 241 GAGATAGGTCAAGTCCAGGGGTTTAAAAATGGGCACCTTCTCTCAGTCCCTCTG 300


```
QY 628 ACACACTACGACAGAGGCACACTAGGCCACTGTCCCACTGTCCCACTCTGTGTGTCTGCC 687
Db 565 -----C 565
QY 688 AGTTCAATTCGAAATGCTCTCGGAGTCAGCAGAGTGGTGCAGACCTGAAAGAACTGAC 747
Db 566 AGTTCAATTCGAAATGCTCTCGGAGTCAGCAGAGTGGTGCAGACCTGAAAGAACTGAC 625
QY 748 CTGCCAGGTGGCTACTCTCAACAACAAAT-----GCCTCACTGAAGGACCTGCTG 798
Db 626 CTGCCAGGTGGCTACTCTCAACAACAAATGGTAGAGAGCTCACTGAAGGACCTGCTG 685
QY 799 CCCCGTCAACTGGGTGGAGACCAAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTC 858
Db 686 CCCCGTCAACTGGGTGGAGACCAAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTC 745
QY 859 CTGGGCGGAGGCTGAGAAGTACTGCCAGCTGAAGAACGCCACCTGGTGGTCACTCAATC 918
Db 746 CTGGGCGGAGGCTGAGAAGTACTGCCAGCTGAAGAACGCCACCTGGTGGTCACTCAATC 805
QY 919 CAGGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGATGGGCT 978
Db 806 CAGGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGATGGGCT 865
QY 979 CAGTGACCTGGAAGGAGCCTGGGAAGTGGGTGGATGGAAACAGACTATGCGACCGGCTTCCA 1038
Db 866 CAGTGACCTGGAAGGAGCCTGGGAAGTGGGTGGATGGAAACAGACTATGCGACCGGCTTCCA 925
QY 1039 GAACCTGGAAGCCAGGCGACGACAGACACTGGCGAGGGGCGAGGCTGGGTGGAGGCGAGGA 1098
Db 926 GAACCTGGAAGCCAGGCGACGACAGACACTGGCGAGGGGCGAGGCTGGGTGGAGGCGAGGA 985
QY 1099 CTGTGCTCACTTCCATCCAGAGCGGAGGTGGAATGACGACGCTGCGACAGGCGCTTACCA 1158
Db 986 CTGTGCTCACTTCCATCCAGAGCGGAGGTGGAATGACGACGCTGCGACAGGCGCTTACCA 1045
QY 1159 CTGGGTCTCGGAGGCTGGGCTCAGACACGACGAGAGTCACTGAGCTGCGCTTTGG 1218
Db 1046 CTGGGTCTCGGAGGCTGGGCTCAGACACGACGAGGAGTCACTGAGCTGCGCTTTGG 1105
QY 1219 TGGGACCAACCGGCGACAGAAATGGCGGTGGGAGGAGGACTCTTCTCAGCAGCTCTCTGC 1278
Db 1106 TGGGACCAACCGGCGCGACAGAAATGGCGGTGGGAGGAGGACTCTTCTCAGCAGCTCTCTGC 1165
QY 1279 AAGACCGCTCTGGGAGAGAAATAAGCACTGGGAGATTGGAAGCACTGCTAACTTTTGA 1338
Db 1166 AAGACCGCTCTGGGAGAGAAATAAGCACTGGGAGATTGGAAGCACTGCTAACTTTTGA 1225
QY 1339 TTTTCTCTTTAAATTTTAAAGATGGTATAGTGTCTTAAGCTTTTATTTTTC 1398
Db 1226 TTTTCTCTTTAAATTTTAAAGATGGTATAGTGTCTTAAGCTTTTATTTTTC 1285
QY 1399 AACTTTTGAAGTCACTTCATGAAGGTATATTTTACATATAAATAATGCACTCATTT 1458
Db 1286 AACTTTTGAAGTCACTTCATGAAGGTATATTTTACATATAAATAATGCACTCATTT 1345
```

RESULT 6

```
US-10-829-107-9
; Sequence 9, Application US/10829107
; Publication No. US20040192892A1
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/10/829,107
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/862,802A
```

```
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1370
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid and protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (273)..(1091)
; OTHER INFORMATION: protein coding sequence
; US-10-829-107-9

Query Match 76.0%; Score 1108; DB 18; Length 1370;
Best Local Similarity 89.4%; Pred. No. 2.6e-297;
Matches 1288; Conservative 0; Mismatches 5; Indels 147; Gaps 3;

QY 28 TAGGGCTCTGGGATTTAGACCCCAAGACCAAGCAGGACTCCAGTCACTTACCCAGCT 87
Db 44 TACTGTCTCTGGATTTTCAGATCCAAGACCAAGCAGGACTCCAGTCACTTACCCAGCT 103
QY 88 CTCAGGACACAGCGCTCCCAACTCTGAGTGACCTCCACCTCTGCTGCTTGCAGACAA 147
Db 104 CTCAGGACACAGCGCTCCCAACTCTGAGTGACCTCCACCTCTGCTGCTTGCAGACAA 163
QY 148 CCAACGTGGGAATCAACCCCTCCAGACCTCCACAGCTCCACCCAGACTGGGCGCCGGC 207
Db 164 CCAACGTGGGAATCAACCCCTCCAGACCTCCACAGCTCCACCCAGACTGGGCGCCGGC 223
QY 208 CTTGCTCTCAATTTAGCTGTGACAACTCAGAGCGTGTGGCCCAAGCATGCAAGGAC 267
Db 224 CTTGCTCTCAATTTAGCTGTGACAACTCAGAGCGTGTGGCCCAAGCATGCAAGGAC 283
QY 268 GTATGAAAACCTCCAGTACTTGGAGAAATAGGTCAAAAGTCCAGGGTTTAAAGATGGCC 327
Db 284 GTATGAAAACCTCCAGTACTTGGAGAAATAGGTCAAAAGTCCAGGGTTTAAAGATGGCC 343
QY 328 ACTTCTCTCCAGTCCCTCTCTGACGCGTCTCCGCTCTGGGCGCTGCCATCTCTGCTGTC 387
Db 344 ACTTCTCTCCAGTCCCTCTCTGACGCGTCTCCGCTCTGGGCGCTGCCATCTCTGCTGTC 362
QY 388 CTGGGCGCTCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 447
Db 363 -----CTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 406
QY 448 ATTTTCAGAGGACCTGGTGAACCTGAGAACAGATTTTAGCAACTTCACTTCAAACTGT 507
Db 407 ATTTTCAGAGGACCTGGTGAACCTGAGAACAGATTTTAGCAACTTCACTTCAAACTGT 466
QY 508 GCGGAGATCCAGGACCTGACTTCCAGGGCAGCAGCTTGGAGAAACATAGCATCTCT 567
Db 467 GCGGAGATCCAGGACCTGACTTCCAGGGCAGCAGCTTGGAGAAACATAGCATCTCT 526
QY 568 GAAAGCTGAGTGGAGGTTTCAAGCAGGAAACGGCAGGAGGTATCTGAGCTCCAGGA 627
Db 527 GAAAGCTGAGTGGAGGTTTCAAGCAGGAAACGGCAGGAGGTATCTGAGCTCCAGGA 564
QY 628 ACACACTAGCAGAGGACACACTAGGCCACTGTCCCACTGTCTGTGTGTCTGCC 687
Db 565 -----C 565
QY 688 AGTTCAATTCGAAATGCTCTCGGAGTCAGCAGAGTGGTGCAGACCTGAAAGAACTGAC 747
Db 566 AGTTCAATTCGAAATGCTCTCGGAGTCAGCAGAGTGGTGCAGACCTGAAAGAACTGAC 625
QY 748 CTGCCAGGTGGCTACTCTCAACAACAAAT-----GCCTCACTGAAGGACCTGCTG 798
```


Db 626 CTGCCAGGTGGCTACTCTCAACACAAATGGTGGAGAGCCTCCACTGAAGGACCTGCTG 685
Qy 799 CCGCGTCAACTGGGTGAGACCAAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTC 858
Db 686 CCGCGTCAACTGGGTGAGACCAAGACAGCTGCTACTGTTCTCTCACTCTGGGATGTC 745
Qy 859 CTGGGCGAGGCTCAGAGTACTGCCAGCTGAAGACGCCACCTGGTGGTCACTCAACTC 918
Db 746 CTGGGCGAGGCTCAGAGTACTGCCAGCTGAAGACGCCACCTGGTGGTCACTCAACTC 805
Qy 919 CAGGAGGAGCAGAAATTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCGCT 978
Db 806 CAGGAGGAGCAGAAATTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCGCT 865
Qy 979 CAGTGACCTCAGAGGAGCCTGGAAGTGGGTGGATGGAACAGACTATCGACCGGCTTCCA 1038
Db 866 CAGTGACCTCAGAGGAGCCTGGAAGTGGGTGGATGGAACAGACTATCGACCGGCTTCCA 925
Qy 1039 GAACCTGAAGCAGGCCAGCCAGACGACTGGCAGGGGCACGGGCTGGGTGGAGCGAGGA 1098
Db 926 GAACCTGAAGCAGGCCAGCCAGACGACTGGCAGGGGCACGGGCTGGGTGGAGCGAGGA 985
Qy 1099 CTGTGCTCACTTCATCCAGACGCGCAGGTGGAATGACGAGCTCTGCCAGAGGCCCTACCA 1158
Db 986 CTGTGCTCACTTCATCCAGACGCGCAGGTGGAATGACGAGCTCTGCCAGAGGCCCTACCA 1045
Qy 1159 CTGGGCTCGAGGCTGGGCTGAGTCCAGACGCGCAGGAGAGTCACTGAGCTGCTTTGG 1218
Db 1046 CTGGGCTCGAGGCTGGGCTGAGTCCAGACGCGCAGGAGAGTCACTGAGCTGCTTTGG 1105
Qy 1219 TGGGACCAACCGGCCACAGAAATGGCGGTGGGAGGAGCTCTCTCAAGACCTCCTCGC 1278
Db 1106 TGGGACCAACCGGCCACAGAAATGGCGGTGGGAGGAGCTCTCTCAAGACCTCCTCGC 1165
Qy 1279 AAGACCCCTCTGGGAGAGAAATAGCACTGGGAGATTGGAGCACTGCTAACTTTTGA 1338
Db 1166 AAGACCCCTCTGGGAGAGAAATAGCACTGGGAGATTGGAGCACTGCTAACTTTTGA 1225
Qy 1339 TTTTCTCTTTAAATTTTAAAGATGGTATAGTCTTTAAAGCTTTTAAATTTTTC 1398
Db 1226 TTTTCTCTTTAAATTTTAAAGATGGTATAGTCTTTAAAGCTTTTAAATTTTTC 1285
Qy 1399 AACTTTTGAAGTCAACTCATGAGGTATAATTTTACATATAAAATGCACTCATTT 1458
Db 1286 AACTTTTGAAGTCAACTCATGAGGTATAATTTTACATATAAAATGCACTCATTT 1345

RESULT 7
US-10-138-588-89
; Sequence 89, Application US/10138588
; Publication No. US20040018594A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: NOVEL ANTIBODIES THAT BIND TO ANTIGENIC POLYPEPTIDES, NUCLEIC ACID
; TITLE OF INVENTION: ENCODING THE ANTIGENS, AND METHODS OF USE
; FILE REFERENCE: 21402-347A
; CURRENT APPLICATION NUMBER: US/10/138,588
; PRIOR FILING DATE: 2002-05-01
; PRIOR FILING DATE: 2001-08-17
; PRIOR FILING DATE: 2001-05-03
; PRIOR FILING DATE: 2001-07-31
; PRIOR FILING DATE: 2001-09-25
; PRIOR FILING DATE: 2001-09-25
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 203
; SEQ ID NO 89
; LENGTH: 1072
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: CDS
; LOCATION: (45) .. (1002)
US-10-138-588-89

Query Match 71.3%; Score 1039.2; DB 17; Length 1072;
Best Local Similarity 98.9%; Pred No. 3.2e-278;
Matches 1060; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

Qy 213 CTCATTTTCAGCTGTGACAACTCAGAGCCGCTGTGTGCCCAAGCATGACAAGGACGTATG 272
Db 1 CTCATTTTCAGCTGTGACAACTCAGAGCCGCTGTGTGCCCTAAGCATGACAAGGACGTATG 60
Qy 273 AAAACTTCCAGTACTTGGAGAAATAAGGTGAAGTCCAGGGGTTTAAAAATGGGCCACTTC 332
Db 61 AAAACTTCCAGTACTTGGAGAAATAAGGTGAAGTCCAGGGGTTTAAAAATGGGCCACTTC 120
Qy 333 CTCTCCAGTCCCTCTCGCAGCGCTCTCCGCTCTGGGCCCTGCCATCTCTGCTGTGCCCTGG 392
Db 121 CTCTCCAGTCCCTCTCGCAGCGCTCTCTGCTCTGGGCCCTGCCATCTCTGCTGTGCCCTGG 180
Qy 393 GCCTCGGCTGTGCTGTGCTGTGCTCATCATCTGTGTGTGATTCACAAATTTCCAAATTTTC 452
Db 181 GCCTCGGCTGTGCTGTGCTGTGCTCATCATCTGTGTGTGATTCACAAATTTCCAAATTTTC 240
Qy 453 AGAGGACCTGTGTGACCTTGAGAAACAGATTTTAGCAACTTCACTCAAACTGTGGCGG 512
Db 241 AGAGGACCTGTGTGACCTTGAGAAACAGATTTTAGCAACTTCACTCAAACTGTGGCGG 300
Qy 513 AGATCCAGGCACTGACTTCCCGAGGCGAGCAGCTTGGAGAAACGATAGCATCTCTGAAG 572
Db 301 AGATCCAGGCACTGACTTCCCGAGGCGAGCAGCTTGGAGAAACGATAGCATCTCTGAAG 360
Qy 573 CTGAGGTGGAGGGTTTCAAGCAGGAACGGCAGGACGGGGTATCTGAGCTCCAGGAACACA 632
Db 361 CTGAGGTGGAGGGTTTCAAGCAGGAACGGCAGGACGGGGTATCTGAGCTCCAGGAACACA 420
Qy 633 CTAGCAGAAAGGACACCTTAGGCCACTGTCCCACTGTGCCCATCTGTGTGTGTCTCCAGTTC 692
Db 421 CTAGCAGAAAGGACACCTTAGGCCACTGTCCCACTGTGCCCATCTGTGTGTGTCTCCAGTTC 480
Qy 693 ATTCTGAATGTCTCTCGGAGTCCAGCAGCTGTGCAAGACCTGGAAGAACTGAAGAACTGACCTGCC 752
Db 481 ATTCTGAATGTCTCTCGGAGTCCAGCAGCTGTGCAAGACCTGGAAGAACTGAAGAACTGACCTGCC 540
Qy 753 AGGTGGTACTCTCAACAAACAAT-----GCCTCCACTGAAGGACCTGTGCCCCG 803
Db 541 AGGTGGTACTCTCAACAAACAATGGTGAAGAGCCTCCACTGAAGGACCTGTGCCCCG 600
Qy 804 TCAACTGGGTGGAGACCAAGACAGCTGTCTACTGTTCTCTCACTCTGGGATGTCTGGG 863
Db 601 TCAACTGGGTGGAGACCAAGACAGCTGTCTACTGTTCTCTCACTCTGGGATGTCTGGG 660
Qy 864 CCAGGCTGAGAGTACTGCCAGCTGAAGAACCCCACTGGTGGTGTCTCACTCACTCCAGGG 923
Db 661 CCAGGCTGAGAGTACTGCCAGCTGAAGAACCCCACTGGTGGTGTCTCACTCACTCCAGGG 720
Qy 924 AGGACAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGTG 983
Db 721 AGGACAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGTG 780
Qy 984 ACCCTGAAGAGCCTGGAGTGGGTGGATGGAACAGACTATGCGACCGGCTTCAGAACT 1043
Db 781 ACCCTGAAGAGCCTGGAGTGGGTGGATGGAACAGACTATGCGACCGGCTTCAGAACT 840
Qy 1044 GGAAGCCAGGCCAGACCACTGGCAGGGGACCGGGCTGGGTGGAGGAGGACTGTG 1103
Db 841 GGAAGCCAGGCCAGACCACTGGCAGGGGACCGGGCTGGGTGGAGGAGGACTGTG 900
Qy 1104 CTCACTTCCATCCAGAGCGGAGGTGGAATGACGACGTCTGCCAGAGGCCCTTACCACTGGG 1163
Db 901 CTCACTTCCATCCAGAGCGGAGGTGGAATGACGACGTCTGCCAGAGGCCCTTACCACTGGG 960

```
QY 1164 TGTGCGAGGCTGGCTGGGTGAGCCAGCCAGGAGAGTCACTGAGCTGCTTTGTGGGA 1223
Db 961 TGTGCGAGGCTGGCTGGGTGAGCCAGCCAGGAGAGTCACTGAGCTGCTTTGTGGGA 1020
QY 1224 CCACCCGGCCACAGAAATGGCGTGGGAGGAGTCTTTCTCAGCACTTCCT 1275
Db 1021 CCACCCGGCCACAGAAATGGCGTGGGAGGAGTCTTTCTCAGCACTTCCT 1072

RESULT 8
US-09-925-302-18
; Sequence 18, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL04
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 929
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (431)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (613)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (918)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (929)
; OTHER INFORMATION: n equals a,t,g, or c
;
US-09-925-302-18
Query Match 49.7%; Score 724.2; DB 9; Length 929;
Best Local Similarity 89.8%; Pred. No. 1.4e-190;
Matches 836; Conservative 6; Mismatches 6; Indels 83; Gaps 3;

QY 1 GTTGAGGAGATGGGATGTCCAGATGATAGGGCTCCTGGGATTTCAAGCCCAAGACCAGC 60
Db 32 GTTGAGGAGATGGGATGTCCAGATGATAGGGCTCCTGGGATTTCAAGCCCAAGACCAGC 91
QY 61 AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGTCCCAACTCTGAGTGAC 120
Db 92 AGGACTCCAGTCACTTACCCAGCTCTCCAGGACACAGCGTCCCAACTCTGAGTGAC 151
QY 121 GTCCCACTCTGTCTTGCAGCACAAACCAAGCTGGGAATCACACCCTCCAGACCTCCCA 180
Db 152 GTCCCACTCTGTCTTGCAGCACAAACCAAGCTGGGAATCACACCCTCCAGACCTCCCA 211
QY 181 CAGCTCCACCCAGACTGGCGCGGCCCTCCCTCCATTTTCAGCTGTGTGACACCTCAGAG 240
Db 212 CAGCTCCACCCAGACTGGCGCGGCCCTCCCTCCATTTTCAGCTGTGTGACACCTCAGAG 271
QY 241 CGGTGTGGCCCAAGCATGACAGGACGTATGAAAACCTTCAGTACTTGGAGAAATAGGT 300
Db 272 CGGTGTGGCCCAAGCATGACAGGACGTATGAAAACCTTCAGTACTTGGAGAAATAGGT 331
QY 301 GAAAGTCCAGGGGTTTAAAAATGGGCCACTTCTCTCCAGTCCCTCCTCAGCGTCTCG 360
Db 332 GAAAGTCCAGGGGTTTAAAAATGGGSCACTTCTCTCCAGTCCCTCCTCAGCGTCTCG 391
QY 361 CTCTGGGCCCCGCCATCTCTGCTGTCTGCTGGGCCCTCGGCCCTGCTGCTGTCATCAT 420
```

```
Db 392 CTCTGGGSCCTGCCATCTCTCTGCTGTCCCTGGGCTCGGNCCTCCTGCTGCTGTCATCAT 451
QY 421 CTGTGTGTTGGATTCCAAAATTTCCAAAATTTCCAGAGGACCTGTGTGACCTTGAGAACAGA 480
Db 452 CTGTGTGTTGGATTCCAAAATTTCCAAAATTTCCAGAGGACCTGTGTGACCTTGAGAACAGA 511
QY 481 TTTTAGCAACTTTCACCTCAACACTGTGTGGCGGAGATCCAGGCACCTGACTTCCAGGGGAG 540
Db 512 TTTTAGCAACTTTCACCTCAACACTGTGTGGCGGAGATCCAGGCA-TGACTTCCAGGGGAG 570
QY 541 CAGCTTGAAGAAACGATAGCATCTCTGAAGAGCTGAGGTGAGGGTTTCAAGCAGGAAACG 600
Db 571 CAGCTTGAAGAAACGATAGCATCTCTGAAGAGCTGAGGTGAGGGTTTCAAGCAGGAAACG 630
QY 601 GCAGCGAGGGGTATCTGAGCTCCAGGAACACACTACCGCAGGAAGGCACACCTAGGCCACTG 660
Db 631 GCAGS----- 635
QY 661 TCCCACTGCCCATCTGTGTGTCTCCAGTTCAATTTCTGAATGTCTCTGCGAGTCCAGCA 720
Db 636 -----CAGTTCAATTTCTGAATGTCTCTGCGAGTCCAGCA 669
QY 721 GCTGTGCAAGACCTGGAAGAACTGACCTGCCAGGTGGCTACTCTCAACAAAT-GCCT 779
Db 670 GCTGTGCAAGACCTGGAAGAACTGACCTGCCAGGTGGCTACTCTCAACAAATGGCCT 729
QY 780 CCAGTGAAGGACCTGTGCTGCCCTCAACTGGGTGGAGCACCAGACAGCTGCTACTTGGT 839
Db 730 CCAGTGAAGGACCTGTGCTGCCCTCAACTGGGTGGAGCACCAGACAGCTGCTACTTGGT 789
QY 840 TCTCTCACTCTGGGATGTCTGGGCGGAGGCTGAGAAAGTACTGCCAGCTGAAGAACGCC 899
Db 790 TCTCTCACTCTGGGATGTCTGGGCGGAGGCTGAGAAAGTACTGCCAGCTGAAGAACGCC 849
QY 900 ACCTGTGTGTCATCAACTCCAGGGAGGAGCA 930
Db 850 ACCTGTGTGTCATCAAAATCCAGGGAGGAGCA 880

RESULT 9
US-09-925-302-18
; Sequence 18, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL04
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 929
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (431)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (613)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (918)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (929)
; OTHER INFORMATION: n equals a,t,g, or c
```

US-09-925-302-18

Query Match 49.7%; Score 724.2; DB 10; Length 929;
Best Local Similarity 89.8%; Pred. No. 1.4e-190;
Matches 836; Conservative 6; Mismatches 6; Indels 83; Gaps 3;

```
QY 1 GTTGAGGAGATGGGATGTCACAGATGATAGGGCTCTGGGATTTCCAGACCCAGACGAGC 60
DB 32 GTTGAGGAGATGGGATGTCACAGATGATAGGGCTCTGGGATTTCCAGACCCAGACGAGC 91
QY 61 AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC 120
DB 92 AGGACTCCAGTCACTCTACCCAGCTCTCCAGGACACAGCGCTCCCAACTCTGAGTGAC 151
QY 121 GTCCCACTCTGTCTCTTGAGACACAAACAAAGTGGGAATCAACACCTCCAGACCTCCCA 180
DB 152 GTCCCACTCTGTCTCTTGAGACACAAACAAAGTGGGAATCAACACCTCCAGACCTCCCA 211
QY 181 CAGTCCACCCAGACCTGGGCGCGGCTCTCCATTTTCAGCTGTGACAACTCAGAG 240
DB 212 CAGTCCACCCAGACCTGGGCGCGGCTCTCCATTTTCAGCTGTGACAACTCAGAG 271
QY 241 CCCTGTTGGCCCAAGCATGACAGGAGTATGAAACTTCCAGTACTTGGAGAAATAGGT 300
DB 272 CCCTGTTGGCCCAAGCATGACAGGAGTATGAAACTTCCAGTACTTGGAGAAATAGGT 331
QY 301 GAAAGTCCAGGGGTTTAAATAATGGGCCACTTCTCTCCAGTCTCTCCAGGCTCTCCG 360
DB 332 GAAAGTCCAGGGGTTTAAATAATGGGCCACTTCTCTCCAGTCTCTCCAGGCTCTCCG 391
QY 361 CTCTGGGCTCTGATCTCTGCTGCTCTGGGCTCTGGGCTCTGCTGCTGCTGCTCAT 420
DB 392 CTCTGGGCTCTGATCTCTGCTGCTCTGGGCTCTGGGCTCTGCTGCTGCTGCTCAT 451
QY 421 CTCTGTTGGTGGATTTCCAAATTTCCAGAGGAGCTTGTGACCTCTGAGAAACAGA 480
DB 452 CTCTGTTGGTGGATTTCCAAATTTCCAGAGGAGCTTGTGACCTCTGAGAAACAGA 511
QY 481 TTTTAGCACTTCACTCAAACTGTGGCGAGATCCAGGCACTGACTTCCAGGCGAG 540
DB 512 TTTTAGCACTTCACTCAAACTGTGGCGAGATCCAGGCACTGACTTCCAGGCGAG 570
QY 541 CAGCTTGGAGAAACGATAGCATCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAAAG 600
DB 571 CAGCTTGGAGAAACGATAGCATCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAAAG 630
QY 601 GCAGGAGGGGTATCTGAGCTCCAGGAAACACATACGACAGAGGACACCTAGGCCACTG 660
DB 631 GCAGG----- 635
QY 661 TCCCCATGCCCATCTGTGTGTGTCTCCAGTTCACTTCTGAAATGCTCTGCGAGTCCAGCA 720
DB 636 -----CAGTTTCACTTCTGAAATGCTCTGCGAGTCCAGCA 669
QY 721 GCTGTTGCAAGACTGAGAAACTGACCTGCCAGGTGGCTACTCTCAACAACTAT-GCCT 779
DB 670 GCTGTTGCAAGACTGAGAAACTGACCTGCCAGGTGGCTACTCTCAACAACTATGGCCT 729
QY 780 CCACTGAAGGACCTGCTGCCCGCTCAACTGGGTGGAGCACCAAGACAGCTGCTACTGGT 839
DB 730 CCACTGAAGGACCTGCTGCCCGCTCAACTGGGTGGAGCACCAAGACAGCTGCTACTGGT 789
QY 840 TCTCTCACTCTGGGATGTCTGGGCGGAGCTGAGAAAGTACTGCCAGCTGAAGAACGCC 899
DB 790 TCTCTCACTCTGGGATGTCTGGGCGGAGCTGAGAAAGTACTGCCAGCTGAAGAACGCC 849
QY 900 ACCTGGTGTCTCACTCACTCCAGGGAGGAGCA 930
DB 850 ACCTGGTGTCTCACTCACTCCAGGGAGGAGCA 880
```

RESULT 10

US-10-138-588-87

; Sequence 87, Application US/10138588
; Publication No. US20040018594A1
; GENERAL INFORMATION:
; APPLICANT: Alsebrook et al.
; TITLE OF INVENTION: NOVEL ANTIBODIES THAT BIND TO ANTIGENIC POLYPEPTIDES, NUCLEIC ACID
; FILE REFERENCE: 21402-347A
; CURRENT APPLICATION NUMBER: US/10/138,588
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/288,395
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/308,901
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/313,388
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/324,757
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/288,900
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 203
; SEQ ID NO 87
; LENGTH: 820
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(804)
US-10-138-588-87

Query Match 49.0%; Score 714.6; DB 17; Length 820;

Best Local Similarity 90.0%; Pred. No. 6.1e-188;

Matches 811; Conservative 0; Mismatches 9; Indels 81; Gaps 1;

```
QY 321 ATGGGGCACTTCTCTCCAGTCCCTCTGAGGCTCTCCGCTCTGGGCTCTGGCCCTGCCATCTCC 380
DB 1 ATGGGGCACTTCTCTCCAGTCCCTCTGAGGCTCTCTGCTCTGGGCTCTGGCCCTGCCATCTCC 60
QY 381 TGTGTCTCTGGGCTCTGGGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 440
DB 61 TGTGTCTCTGGGCTCTGGGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 441 ATTCCAAATTTTCAGAGGACCTGCTGACCTCTGAGAACAGATTTTAGCACTTCACTCTCAA 500
DB 121 ATTCCAAATTTTCAGAGGACCTGCTGACCTCTGAGAACAGATTTTAGCACTTCACTCTCAA 180
QY 501 ACATGTGGCGAGATCCAGGCACTGACTTCCAGGCGAGCAGCTTGGAGAAACGATAG 560
DB 181 ACATGTGGCGAGATCCAGGCACTGACTTCCAGGCGAGCAGCTTGGAGAAACGATAG 240
QY 561 CATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAAACGGCAGG----- 285
DB 241 CATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAAACGGCAGG----- 285
QY 621 TCCAGGACACACTACGACAGGACACCTAGGCCACTGTCCCCTCTGCTGTGTGTGTGTGTGT 680
DB 286 ----- 285
QY 681 GTGTCCAGTTCACTTCTGAAATGCTCTGCGAGTCCAGCAGCTGTTGAAAGACCTGAAGA 740
DB 286 -----CAGTTTCACTTCTGAAATGCTCTGCGAGTCCAGCAGCTGTTGAAAGACCTGAAGA 339
QY 741 AACTGACCTGCCAGGTGGTGTCTCTCAAACAAATGCTCTCACTGAAAGGACCTGTGTGCC 800
DB 340 AACTGACCTGCCAGGTGGTGTCTCTCAAACAAATGCTCTCACTGAAAGGACCTGTGTGCC 399
QY 801 CCCTCACTGGGTGGAGCACCAAGACAGCTGTACTGTTTCTCTCACTCTGGGATGTCTCT 860
DB 400 CTCTCACTGGGTGGAGCACCAAGACAGCTGTACTGTTTCTCTCACTCTGGGATGTCTCT 459
QY 861 GGGCCGAGGCTGAGAAAGTACTGCGAGCTGAAGAACGCCACCTGGTGTGTGTGTGTGTGTGT 920
DB 460 GGGCCGAGGCTGAGAAAGTACTGCGAGCTGAAGAACGCCACCTGGTGTGTGTGTGTGTGTGT 519
```

QY 921 GGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCCATACACCTGGATGGGCTCA 980
Db 520 GGGAGGAGCAGAAATTTTGTCCAGAAATATCTAGGCTCCCATACACCTGGATGGGCTCA 579
QY 981 GTGACCTTGAAGGAGCCTGGAAGTGGGTGGATGGGAACAGACTATGGACCGGCTTCCAGA 1040
Db 580 GTGACCTTGAAGGAGCCTGGAAGTGGGTGGATGGGAACAGACTATGGACCGGCTTCCAGA 639
QY 1041 ACTGGAAGCCAGGCCAGCAGCAGCTGGCAGGGGCAACGGGCTGGGTGGAGCGGAGGACT 1100
Db 640 ACTGGAAGCCAGGCCAGCAGCAGCTGGCAGGGGCAACGGGCTGGGTGGAGCGGAGGACT 699
QY 1101 GTGCTCACTTCCATCCAGACGGCAGGTGGAATGACGACGCTCTGCCAGAGGCCCTTACCAC 1160
Db 700 GTGCTCACTTCCATCCAGTCCGCGAGGTGGAATGACGACGCTCTGCCAGAGGCCCTTACCAC 759
QY 1161 GGGTCTGCGAGGCTGGCTGGTCCAGACCGAGGAGAGTCACTGAGGTGCTGGTGGTG 1220
Db 760 GGGTCTGCGAGGCTGGCTGGTCCAGACCGAGGAGAGTCACTGAGGTGCTGGTGGTG 819
QY 1221 G 1221
Db 820 G 820

RESULT 11
US-10-152-319A-1777
; Sequence 1777, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgins, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashof, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1777
; LENGTH: 1358
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_022393
US-10-152-319A-1777

Query Match 30.0%; Score 438; DB 17; Length 1358;
Best Local Similarity 68.1%; Pred. No. 7.9e-111;

Matches 676; Conservative 0; Mismatches 295; Indels 21; Gaps 4;
QY 210 TGCCTCCAAATTTTTCAGCTGTGCAACCTCAGAGCC---GTGTTGGCCCAAGCATGACAAAGA 266
Db 122 TGTCTCGGTTTCAGTTTCAGATAGCCCTTAGAGCCCTTGTGTTTGTGCAAGCATGACATGG 181
QY 267 CGTATGAAACTTCCAGTACTTGGAGAAATAGAGTGAAGTCCAGGGGTTTAAAAATGGGC 326
Db 182 CATATGAAACTTCCAGAACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 239
QY 327 CACTTCTCTCAGTCCCTCTGAGCGTCTCCGCTCTGGGCCCTTGGCATCTCTCTGCTGT 386
Db 240 ----TCCCTCCCACTCCCTCTGCAATATCTCTCTGAGACCCCTCTCTCTCTCTCT 295
QY 387 CCCTGGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 446
Db 296 CCCTGGGCTTAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 355
QY 447 AATTTTCAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 506
Db 356 AGTTAAGGAGGAGCTTAGAAACCTGAGAACCACTTTTAGACAAACACCACTTCAACACCA 415
QY 507 TGGCGGAGATCCAGGCACTGATTTCCAGGCGCAGGCTTGGAGAGAAACGATAGCATCTC 566
Db 416 AGGCTGAACTACAGGCGCTTGGCTTCCAGGGGTGACAGCTTGCNAAACAGGAATCAATTCT 475
QY 567 TGAAGCTGAGCTGGAGGGTTTCAAGCAGAGAAACGCGCAGGAGGAGGCTATCTGAGCTC 626
Db 476 TGAAGTGGAGTGGATGATCATGGGCAGGAACTTGCAGGCGAGGCGGAGGCTTGGAGC 535
QY 627 AACACACTACGAGAGGACACCTAGGCCACTGTGCCCACTGTGCCCACTGTGCTGTGTG 686
Db 536 AGGTGGCTTCTTGGAGAGCACAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 586
QY 687 CAGTTCACTTGAATGCTCTCTGAGCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
Db 587 CAGATCTATCTGAAATTAACCGATCGTGTGCAACAGCTGGGAGGAGGAGGAGGAGGAG 646
QY 747 CTGCGAGGTGGCTACTCTCAACCAACATGCTCCACTGAAAGGAGGAGGAGGAGGAGGAG 806
Db 647 CATGTGAGTGGCGAGCTTCAAGAAACAGGCTGAGC---AGTGGCTGTGCTGCTGCTGCT 703
QY 807 ACTGGTGGAGCACCAAGACAGCTGCTACTGTGCTTCTCTCACTTGGGATGCTCTGGGCG 866
Db 704 ACTGGATGGAGCATGAAGGCGAGCTGCTACTGTGCTTCTCTCACTTGGGAGGAGGAGG 763
QY 867 AGGCTGAGAGTACTGCGAGCTGAAGACGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
Db 764 AAGCTGACACAGTACTGCGAGCTGGAGAAATTCNACCTAGTGGTGGTCAACTCCCTGG 823
QY 927 AGCAGAAATTTTGTCCAGAAATATCTAGGCTCCGCATACACCTGGATGGGCTCAGTG 986
Db 824 AGCAGAAATTTTCTACAGACTCAGATGGGAGTGGTCACTTGGATGGGCTTACCGAGC 883
QY 987 CTGAAGAGGAGCTTGAAGTGGGTGGATGGAAACAGACTATGCGACCGGCTTCCAGAACT 1046
Db 884 AAAATGGGCTTGGCGATGGGTGGAGCGGACCGACTATGAGAAAGGCTTTACGCACTGG 943
QY 1047 AGCCAGGCGCAGCAGACGACTGGCAGGGGCGCGGCTGGGTGGGAGGAGGAGGAGGAG 1106
Db 944 CCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1003
QY 1107 ACTTCCATCCAGACGGCAGGTGGAAATGACAGCTGTGCCAGAGGCGCTTACCACCTGG 1166
Db 1004 ACTTCACTCAGATGGTCTGCTGGAATGATGAGCTTTGCCAGAGGCGCTTACCGCTGG 1063
QY 1167 GCGAGGCTGGCTGGGTGAGACCGAGGAGGAG 1198
Db 1064 GTGAGATGAAGCTGGCCAAAGGAGCAGCTAGGAG 1095

RESULT 12
US-09-880-107-2230

; Sequence 2230, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2230
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M10058
US-09-880-107-2230

Query Match 26.2%; Score 382; DB 9; Length 1277;
Best Local Similarity 66.1%; Pred. No. 3e-95;
Matches 646; Conservative 0; Mismatches 245; Indels 87; Gaps 3;

QY 224 CTGTGACCACTCAGAGCCGTGTGGCCCAAGCATGACAGGACGATGAAATTCAG 283
DB 140 CTGAGCAATCCCAGGTCAGCGCCCTATCATGCAAGGAGTATCAAGACCTTCAG 199

QY 284 TACTTGAGAAATAGGTGAAAGTCCAG--GGGTTTAAATAATGGCCACTTCCTCTCCAG 340
DB 200 CATCTGACATGAGAGAGTACACCATCAGCTCAGAAAGGAGGACCTCTCCAG 259

QY 341 TCCTCTCTGAGCGTCTCCGCTCTGGGCGCTTGCATCTCTCTGCTCTCTCTGGGCTCGGC 400
DB 260 CCCCTCTGAGCGTCTCTGCTCCGACCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 319

QY 401 CTGCTCTCTGCT 460
DB 320 CT 379

QY 461 CTGCTGACCTCAGAGACAGATTTAGCACTTCACTTCAAACTCAAACTCAAACTCAAACTCA 520
DB 380 CTGCGGGCTCAGAGAGAGTTCAGCACTTACAGCGAGCAGCGAGGCGGAGGAG 439

QY 521 GCCTGACCTCAGAGACAGATTTAGCACTTCACTTCAAACTCAAACTCAAACTCAAACTCA 580
DB 440 GGCTTGAGCACCAGGAGGCAATGTGGGAAGAAAGATGAAGTCGCTAGAGTCCAGCTG 499

QY 581 GAGGTTTCAAGCAGGACCGGAGGAGGAGGATCTGAGCTCCAGGACACACTACGAC 640
DB 500 GAGAAACAGCAGAGGACCTGAGT-----524

QY 641 AAGGCACACCTAGGCCACTGTCCCACTGTCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 700
DB 525 -----AAGTCACTCCAGC 538

QY 701 ATGCTCTCTGAGTCCAGAGCTGGTCAAGACTGAGAACTGAGAACTGAGCTCCAGAGTGGCT 760
DB 539 CTGCTCTCTCAGTGAAGCAGTTCGTCTGACCTGGGAGCCTGAGCTGTGAGATGGCG 598

QY 761 ACTCTCAACAAATGCTCTCACTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 820
DB 599 GCGCTCCAGGCAATGGCT---CAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 655

QY 821 CAAGACAGGCTGCT 880
DB 656 GAGGCGAGCTGCT 715

QY 881 TGCAGCTGAAGAACGCCACCTGGTGGTTCATCACTCAGGAGGAGGAGGAGGAGGAGGAGGAG 940
DB 716 TGCAGCTGAAGAACGCCACCTGGTGGTTCATCACTCAGGAGGAGGAGGAGGAGGAGGAGGAG 775

QY 941 CAGAAATATCTAGGCTCCGATACACTGATGGGCTCAGTCACTGAGGAGGAGGAGGAGGAGGAG 1000
DB 776 CAGCAACATAGGCTCTGTAACACTGATGGGCTCAGCAACCAAAAGGAGGAGGAGGAGGAG 835

QY 1001 AAGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1060
DB 836 AAGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 895

QY 1061 GAGCACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1120
DB 896 GAGCACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 955

QY 1121 GGCAGGTGGAATGACGACGCTCTGCCAGAGGCGCTTACCACTGGGCTCTGCCAGGCTGGGCTG 1180
DB 956 GGCAGGTGGAATGACGACGCTCTGCCAGAGGCGCTTACCACTGGGCTCTGCCAGAGGAGGAG 1015

QY 1181 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1198
DB 1016 GACAAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1033

RESULT 13
US-10-305-720-1186
; Sequence 1186, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1186
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g179078
US-10-305-720-1186

Query Match 26.2%; Score 382; DB 17; Length 1277;
Best Local Similarity 66.1%; Pred. No. 3e-95;
Matches 646; Conservative 0; Mismatches 245; Indels 87; Gaps 3;

QY 224 CTGTGACCACTCAGAGCCGTGTGGCCCAAGCATGACAGGACGATGAAATTCAG 283
DB 140 CTGAGCAATCCCAGGTCAGCGCCCTATCATGCAAGGAGTATCAAGACCTTCAG 199

QY 284 TACTTGAGAAATAGGTGAAAGTCCAG--GGGTTTAAATAATGGCCACTTCCTCTCCAG 340
DB 200 CATCTGACATGAGAGAGTACACCATCAGCTCAGAAAGGAGGACCTCTCCAG 259

QY 341 TCCTCTCTGAGCGTCTCCGCTCTGGGCGCTTGCATCTCTCTGCTCTCTCTGGGCTCGGC 400
DB 260 CCCCTCTGAGCGTCTCTGCTCCGACCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 319

QY 401 CTGCTCTCTGCT 460
DB 320 CT 379

QY 461 CTGCTGACCTCAGAGACAGATTTAGCACTTCACTTCAAACTCAAACTCAAACTCAAACTCA 520
DB 380 CTGCGGGCTCAGAGAGAGTTCAGCACTTACAGCGAGCAGCGAGGCGGAGGAG 439

QY 521 GCCTGACCTCAGAGACAGATTTAGCACTTCACTTCAAACTCAAACTCAAACTCAAACTCA 580
DB 440 GGCTTGAGCACCAGGAGGCAATGTGGGAAGAAAGATGAAGTCGCTAGAGTCCAGCTG 499

QY 581 GAGGTTTCAAGCAGGACCGGAGGAGGAGGATCTGAGCTCCAGGACACACTACGAC 640
DB 500 GAGAAACAGCAGAGGACCTGAGT-----524

QY 641 AAGGCACACCTAGGCCACTGTCCCACTGTCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 700
DB 525 -----AAGTCACTCCAGC 538

QY 701 ATGCTCTCTGAGTCCAGAGCTGGTCAAGACTGAGAACTGAGAACTGAGCTCCAGAGTGGCT 760
DB 539 CTGCTCTCTCAGTGAAGCAGTTCGTCTGACCTGGGAGCCTGAGCTGTGAGATGGCG 598

QY 761 ACTCTCAACAAATGCTCTCACTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 820
DB 599 GCGCTCCAGGCAATGGCT---CAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 655

QY 821 CAAGACAGGCTGCT 880
DB 656 GAGGCGAGCTGCT 715

```
Db 440 GGCTTGAGCAGCCAGGAGGCAATGTGGGAAGAAAGATGAAGTCGCTAGTCCAGCTG 499
Qy 581 GAGGGTTTCAAGCAGGAACGGCAGGAGGATCTGAGCTCCAGGACACACTACGAG 640
Db 500 GAGAAACAGCAGAGACCTGAGTG-----524
Qy 641 AAGGCACACTAGGCCACTGTGCCCACTGCCCATCTGTGTGTGCCAGTTCATTCTGAA 700
Db 525 -----NAGATCACTCCAGC 538
Qy 701 ATGCTCTCGAGTCAGCAGCTGTGGTGAAGACCTGAAGAACTGACCTGCCAGTGGCT 760
Db 539 CTGCTCTCCAGCTGAAGCAGTTCGTGTCTGACCTGCGGAGCTGAGCTGTGAGATGGC 598
Qy 761 ACTCTCAACAACATGCTCCACTGAAGGAGCTGCTGCCCGCTCAACTGGGTGGAGC 820
Db 599 GCGCTCCAGGGCAATGGCT---CAGAAAGAGCTGTGCTGCCGGTCAACTGGGTGGAGC 655
Qy 821 CAAGCAGCTGTACTGTGTTCTCTCACTCTGGGATGTCTCTGGCGGAGGCTGAGAGTAC 880
Db 656 GAGCGAGCTGTACTGTGTTCTCTCGCTCGGGNAGGCTTGGCTGACGCCGACAACTAC 715
Qy 881 TGCAGCTGAAGAACGCCCACTGTGTGTCATCAACTCCAGGAGGAGCAGAAATTTGTG 940
Db 716 TGCCTGCTGGAGAGCGCAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 775
Qy 941 CAGAAATATCTAGGCTCCGCTACATACCTGATGGGCTTCACTGACCTGAAAGGAGCTGG 1000
Db 776 CAGCACCACATAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 835
Qy 1001 AAGTGGGTGGATGGAACAGACTATGCGACCGGCTTCCAGAACTGGAAGCCAGCCAGCA 1060
Db 836 AAGTGGGTGGACGGAGCAGCTACGAGACGGCTTCAAGAACTGAGGCGGAGCAGCG 895
Qy 1061 GACGACTGCGAGGGGCAACGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1120
Db 896 GACGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 955
Qy 1121 GCGAGTGGTAATGACAGCTCTGCGAGAGGCGCTTACCACTGGTGTGAGGCTGCGCTG 1180
Db 956 GCGCGTGTGAAGACAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1015
Qy 1181 GGTGACAGCAGCCAGGAG 1198
Db 1016 GACAGGCCAGCCAGGAG 1033
```

RESULT 14

```
US-10-236-392-9
; Sequence 9, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladimir
; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRochelle, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
```

```
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard A
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-442A
; CURRENT APPLICATION NUMBER: US/10/236,392
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US60/390,155
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US09/635,949
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US60/318,765
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/357,303
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US60/367,753
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US60/369,479
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US09/659,634
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: US60/318,120
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US60/318,130
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 794
; SOFTWARE: Custom
; SEQ ID NO 9
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(982)
US-10-236-392-9
```

Query Match 22.0%; Score 321.4; DB 17; Length 1055;

Best Local Similarity 62.1%; Pred. No. 2e-78;

Matches 591; Conservative 0; Mismatches 276; Indels 84; Gaps 2;

```
Qy 248 GGCCCAAGCATGACAAAGACGATGAAACCTTCCAGTACTTGGAGATAAGGTGAAAGTC 307
Db 112 GGCCCCATCATGGCCAGGACTTTCAGATATCCAGAGCTGAGCTGGAGAGAAATGAC 171
Qy 308 CAGGGGTTTAAATAATGGGCCACTTCTCTTCCAGTCCCTCTGCGAGCGGTCTCCGCTCTGG 367
Db 172 CATCTTTTCCATCAAGGGCCACCTCTGCCCCAGCCCTGGCACAGCGTCTCTGCTCCATG 231
Qy 368 CCTTGCCATCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
Db 232 GTCTGCTTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 291
Qy 428 GTTGATTTCAAAATTTCCAAAATTTTCAGAGGAGCTGCTGAGCCCTGAGAAACAGATTTTAC 487
Db 292 ACTGGTCCCAAGATGACAGCTGCAAGCCAGCTGCGAGCTTGAAGAGCTTTTCAGC 351
Qy 488 AACTTCACCTCAAAACACTGTGGCGAGATCCAGGACCTGACTTCCCGAGGAGCAGCTTG 547
Db 352 AACTTCTCTCGAGCACCCTGACGGAGGTCCAGGCAATCAGCAATCCAGGAGGAGCGCTG 411
Qy 548 GAAGAAACGATAGCATCTCTGAAAGCTGAGTGGAGGGTTTCAAGCAGGAGGAGCGCAG 607
Db 412 GGTGACAAAGATCACATCCCTAGGAGCCAAAGCTGAGGAGAAACAGCAGCAGGAGCTGAAAG-- 469
```

```
QY 608 GGGGTATCTGAGCTCCAGGAACACACTACGAGAGGACACACCTAGGCCACTGTCCCCAC 667
Db 470 -----469
QY 668 TGCCCATCTGTGTGTCTCCAGTTTCATCTCTGAATGCTCTCTGGAGTCCAGCAGCTGGTG 727
Db 470 -----CAGATCAGATGCCCTGTCTTCATCTGAAGCATTCCCC 510
QY 728 CAAGACCTGAAGAACTGACTGCGAGGTGGCTACTCTCAACAACTATGCTCCACTGAA 787
Db 511 GTGGACCTGGCTGTCTGGCTGTCCAGATGAGCTCTCCACAGCAACGGCTCC---CAA 567
QY 788 GGGACCTGCTGCCCGTCACTGGGTGGAGCACCAAGACAGCTGCTACTGTCTCTCAC 847
Db 568 AGGACCTGCTGCCCGTCACTGGGTGGAGCACCAAGCAGCTGCTACTGTCTCTCAC 627
QY 848 TCTGGGATGTCTGGGCGGAGGCTGAGAAGTACTGCGCAGCTGAGAAGACGCCCTGGTG 907
Db 628 TCCGGNAGGCTGGGCTGAGGCGGAGAGTACTGCGCAGCTGGNAGACGACACCTGGTG 687
QY 908 GTCAATCACTCCAGGAGGAGAGAAATTTGTCCAGAAATATCTAGGCTCCGATACAC 967
Db 688 GTCAATCACTCCAGGAGGAGAGAAATTTGTCCAGAAATATCTAGGCTCCGATACAC 747
QY 968 TGGATGGCTCTGAGTACCTGAGGAGCTGGAAGCTGGAGTGGGTGGATGGAACAGACTATGG 1027
Db 748 TGGATAGGCTCTCAGGACAGTGTGGCTCTTTGGAAATGGGTGGATGGACAGACTATAGG 807
QY 1028 ACCGGCTTCCAGAACTCGAAGCCAGGCGCAGCAGCACTGCGAGGGGCGACGGGCTGGGT 1087
Db 808 CACAACTACAGAACTGGGCTGTCTCAGCAGCAGATTAATTTGGCAGGGGCGACGAGCTGGT 867
QY 1088 GGAGGCGAGGAGTGTCTCTCACTTCCATCCAGACGGCAGGTGGAAATGACAGCTGTGCCAG 1147
Db 868 GGAAGTGAAGACTGTGTGAAGTCCAGCCGATGGCGCTGGAACGATGACTTCTGCTGT 927
QY 1148 AGGCCCTACACTGGGTCTGGAGGCTGGCTGGGTGAGACAGCAGCAGGAG 1198
Db 928 CAGGTGTACCGCTGGGTGTGTGAGAAAGCGGAATGCCACCGCGAGGTG 978
```

RESULT 15

```
US-09-880-107-3731
; Sequence 3731, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3731
; LENGTH: 1300
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X55283
US-09-880-107-3731
```

```
Query Match 22.0%; Score 321.4; DB 9; Length 1300;
Best Local Similarity 62.1%; Pred. No. 2.2e-78;
Matches 591; Conservative 0; Mismatches 276; Indels 84; Gaps 2;
```

```
QY 248 GGGCCAGCATGACAAGGAGCTATGAAACTTCCAGTACTTTGGAGTAATAGGTGAAGTC 307
Db 182 GGGCCCATCATGCGCCAGGACTTTCAAGATATCAGCAGCTGAGCTCGGAGGAAATGAC 241
QY 308 CAGGGGTTTTAAATATGGGCACATTCCTCTCCAGTCCCTCTCGAGCGCTCTCCGCTCTGGG 367
Db 242 CATCTTTCCATCAAGGGCCACCTCTCTGCCAGCCCTGGSCACAGCGTCTCTGCTCCATG 301
QY 368 CCTGCGATCTCTGCTGTCTCTGGGCTCTGGGCTGTGCTGTGCTGTGCTCATCTGTGTG 427
Db 302 GTCTGCTCTGAGTGTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 361
QY 428 GTTGGATTTCCAAATTTCCAAATTTCCAAATTTCCAAATTTCCAAATTTCCAAATTTTTC 487
Db 362 ACTGGGTCCCAAGTGCACAGCTGCAAGCCAGCTGCGAGAGCTTGAAGAGCTTTTCAGC 421
QY 488 AACTTTCACCTTCAAACTACTGTGGCGGAGATCCAGGCACTGACTTTCCAGGGCAGCAGTTG 547
Db 422 AACTTCTCTCGAGCACCTCTGACGGAGTCCAGGCAATCAGCACCCCAACGAGGCGAGGTG 481
QY 548 GAAGAAACGATAGCATCTCTGAAAGCTGAGGTGGAGGTTTCAAGCAGGAAACGGCAGGCA 607
Db 482 GGTGCAAGATCATCTCTTAGGAGCCAGCTGGAGAAACAGCAGCAGGACCTGAAAG-- 539
QY 608 GGGGTATCTGAGCTCCAGGAAACACTACGACAGAGCACACTAGGCCACCTGTGCCCCAC 667
Db 540 -----539
QY 668 TGCCCATCTGTGTGTGTCTCCAGTTTCATCTTGAATGCTCTCTGGAGTCCAGCAGCTGGTG 727
Db 540 -----CAGATCAGATGCCCTGTCTTCCATCTGAAGCACTTCCCC 580
QY 728 CAAGACTTGAAGAACTGACTTCCAGGTGGCTACTTCAACAACTATGCTTCCACTGAA 787
Db 581 GTGACCTTGGCTTCTGGCTTCCAGATGAGGTCTCTCCACAGCAACGGCTCC---CAA 637
QY 788 GGGACCTGTGCGCCGCTCACTGGGTGGAGCACCAAGACAGCTGCTACTGTCTCTCAC 847
Db 638 AGGACCTGTGCGCCCGCTCACTGGGTGGAGCACCAAGGAGCTGCTACTGTCTCTCAC 697
QY 848 TCTGGGATGTCTGGGCGGAGGCTGAGAAGTACTTGGCAGCTGAGAAGACGCCCTGGTG 907
Db 698 TCCGGNAGGCTGGGCTGAGGCGGAGAGTACTTCCAGCTGGAGAACGACACACTGGTG 757
QY 908 GTCAATCACTCCAGGAGGAGAGAAATTTGTCCAGAAATATCTAGGCTCCGATACAC 967
Db 758 GTCAATCACTCTCGGAGGAGCAGAAATTTCAATGTAACAACACAGCAACCCCTTCAATACC 817
QY 968 TGGATGGCTCTGAGTACCTGAGGAGCTGGAAGTGGGTGGATGGAACAGACTATGCG 1027
Db 818 TGGATAGGCTCTCAGGACAGTGTGCTTGGAAATGGGTGGATGGCAGACTATAGG 877
QY 1028 ACCGGCTTCCAGAACTGGAAGCCAGGCGCAGCAGCACTGCGAGGGGCGCAGGCTGGGT 1087
Db 878 CACAACTACAAAGAACTGGGCTGTCTCACTCAGCCAGATTAATTTGGCAGGGCAGGCTGGGT 937
QY 1088 GGAGGCGAGGAGTGTGCTCTCACTTCCATCCAGACGGCAGGTGGAAATGACAGCTGTGCCAG 1147
Db 938 GGAAGTGAAGACTGTGTGAAGTCCAGCCGAGTGGCGCTGGAACGATGACTTCTGCTGTG 997
QY 1148 AGGCCCTACCACTGGGTCTCGGAGGCTGGCTGGGTGAGACAGCAGCAGGAG 1198
Db 998 CAGGTGTACCGCTGGGTGTGTGAGAAAGCGGAATGCCACCGCGAGGTG 1048
```

Search completed: May 28, 2005, 02:07:55
Job time : 928 secs

This page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 12:44:45 ; Search time 65 Seconds
(without alignments)
2489.493 Million cell updates/sec

Title: US-10-829-107-4

Perfect score: 1738

Sequence: 1 MPTYENFQYLEKNKVKVGF.....QRPYHWCAGLGQTSQESH 316

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1730	99.5	316	2	Q8IUN9
2	1544	88.8	292	2	Q14538
3	998.5	57.8	256	2	Q6PIW3
4	918.5	52.8	290	1	LECH HUMAN
5	887	51.0	306	1	MMGL RAT
6	877	50.5	284	2	Q91V84
7	874	50.3	283	1	LECH RAT
8	870	50.1	332	2	Q8JZNI
9	866	49.8	283	1	LECH MOUSE
10	805	46.3	304	2	Q91VT3
11	804	46.3	304	1	MMGL MOUSE
12	791	45.5	361	2	Q8BUD5
13	767	44.1	306	2	Q7Z4G9
14	754.5	43.4	311	1	LECI HUMAN
15	671.5	38.6	301	1	LECI MOUSE
16	633.5	36.4	301	1	LECI RAT
17	619.5	35.6	142	2	Q6DUK3
18	363	20.9	144	2	Q6DD06
19	344	19.8	98	2	Q6DUK2
20	331.5	19.1	381	1	C209_HYLSY
21	326.5	18.8	550	1	KUCR RAT
22	325.5	18.7	381	1	C209_MAGNE
23	325.5	18.6	404	1	C209_PONPY
24	322	18.5	399	1	209L_HYLCO
25	320.5	18.4	381	1	C209_CERAE
26	320.5	18.4	422	1	209L_HYLSY
27	320	18.4	404	1	C209_MACMU
28	319	18.4	358	1	C209_PAPHA
29	319	18.4	399	1	209L_HYLLA
30	316.5	18.2	353	2	Q69F40
31	316.5	18.2	376	1	209L_GORGO

32	316	18.2	404	1	C209_HUMAN
33	313.5	18.0	427	1	C209_GORGO
34	313	18.0	404	1	C209_HYLCO
35	313	18.0	742	2	Q6P9F2
36	313	18.0	742	2	Q9BYH7
37	312	18.0	399	1	209L_HUMAN
38	312	18.0	427	1	C209_PANTR
39	311.5	17.9	548	2	Q8BLZ8
40	311	17.9	417	2	Q8TCR2
41	311	17.9	742	2	Q8WZA4
42	310.5	17.9	450	1	C209_HYLLA
43	309.5	17.8	548	1	KUCR_MOUSE
44	308	17.7	445	1	209L_PANTR
45	304.5	17.5	260	2	Q7TSU6

Q8nx6 homo sapien
Q8hx8 gorilla gor
Q8hy01 hylobates c
Q8p9f2 homo sapien
Q8byh7 homo sapien
Q8h2x3 homo sapien
Q8hxz7 pan troglod
Q8blz8 mus musculu
Q8tc22 homo sapien
Q8wza4 homo sapien
Q8hy03 hylobates l
P70194 mus musculu
Q8hy0 pan troglod
Q7tsu6 mus musculu

ALIGNMENTS

RESULT 1

Q8IUN9	ID	Q8IUN9	PRELIMINARY;	PRT;	316 AA.
AC	Q8IUN9;				
DT	01-MAR-2003	(T-EMBLrel. 23, Created)			
DT	01-MAR-2003	(T-EMBLrel. 23, Last sequence update)			
DT	01-MAR-2004	(T-EMBLrel. 26, Last annotation update)			
DE	C-type lectin, superfamily member 13, isoform 1.				
GN	Name=CLECSF14;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RX	MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,				
RA	Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,				
RA	Jones S.J., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RL	and mouse cDNA sequences";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RP	[2]				
RC	SEQUENCE FROM N.A.				
RN	TISSUE=Brain;				
RA	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.				
RA	EMBL; BC039011; AAH3011.1; -				
DR	HSP; P07306; IDV8.				
DR	Genew; HGNC:16916; CLECSF14.				
DR	GO; GO:0016020; C-membrane; IEA.				
DR	GO; GO:0005529; F-sugar binding; IEA.				
DR	InterPro; IPR002353; Antifreeze1.				
DR	InterPro; IPR001304; Lectin_C.				
DR	InterPro; IPR005640; Lectin_N.				
DR	Pfam; PF00059; Lectin_C; 1.				
DR	Pfam; PF03954; Lectin_N; 1.				
DR	PRINTS; PR00356; ANTIFREEZE1.				
DR	SMART; SM00034; CLECT; 1.				

```

DR PROSITE; P800615; C_TYPE LECTIN_1; 1.
DR PROSITE; P850041; C_TYPE LECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 316 AA; 35446 MW; D3B7193E2E1F58AF CRC64;

Query Match          99.5%; Score 1730; DB 2; Length 316;
Best Local Similarity 99.7%; Pred. No. 2.2e-136;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLRSGPCHLLSLGLGLLLVLCVVG 60
DB 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLRSGPCHLLSLGLGLLLVLCVVG 60

QY 61 QNSKFQDLVLTDRDFTSNFTNTVAEIQALTSQSSLEETIASLKAEEVGFQKQRAQVS 120
DB 61 QNSKFQDLVLTDRDFTSNFTNTVAEIQALTSQSSLEETIASLKAEEVGFQKQRAQVS 120

QY 121 ELQEHHTQKAHLGHCHPCPSVCPVHSEMLLRVQQLVQDLKKLTQCVATLNNN---ASTE 177
DB 118 -----VHSEMLLRVQQLVQDLKKLTQCVATLNNGGEASTE 153

QY 178 GTCCPVNWHVHODSCYWFSGMSWAEEKYCOLKNAHLVWINSREONFVKYLGSAYT 237
DB 154 GTCCPVNWHVHODSCYWFSGMSWAEEKYCOLKNAHLVWINSREONFVKYLGSAYT 213

QY 238 WMGLSDPEGAKWKVDGTDYATGFQNKWKPQDDWQHGGLGGEDCAHFHPDGRWDDVCQ 297
DB 214 WMGLSDPEGAKWKVDGTDYATGFQNKWKPQDDWQHGGLGGEDCAHFHPDGRWDDVCQ 273

QY 298 RPYHWVCEAGLGQTSQESH 316
DB 274 RPYHWVCEAGLGQTSQESH 292

RESULT 3
Q6PIW3 PRELIMINARY; PRT; 256 AA.
ID AC Q6PIW3
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE CLCSF14 protein.
GN Name=CLECSF14;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027858; AAH27858.1; -.
DR GO; GO:0016030; C:membrane; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.

```

DR InterPro: IPR005640; lectin_N.
 DR Pfam: PF00059; Lectin_C; 1.
 DR Pfam: PF03954; Lectin_N; 1.
 SQ SEQUENCE 256 AA; 28958 MW; C2C1A576F369647D CRC64;
 Query Match 57.5%; Score 998.5; DB 2; Length 256;
 Best Local Similarity 87.6%; Pred. No. 2.7e-75;
 Matches 197; Conservative 1; Mismatches 0; Indels 27; Gaps 1;
 QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSGPCHLLSLGLGLLLVVICVVG 60
 DB 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSGPCHLLSLGLGLLLVVICVVG 60
 QY 61 QNSKFORDLVTLTDFSNFTNTVAETQALTSQSSLEETIASLKAEVEGFKERQAGVS 120
 DB 61 QNSKFORDLVTLTDFSNFTNTVAETQALTSQSSLEETIASLKAEVEGFKERQAGVS 117
 QY 121 ELQEHTTQKAHLGHCPCPSVCPVHSEMLLRVQQLVQDLKLTTCVATLNNASTEGTC 180
 DB 118 -----VHSEMLLRVQQLVQDLKLTTCVATLNNASTEGTC 153
 QY 181 CPVNVWEHQDSCYFWSHSGMSWAEAEKYCOLKNAHLVWVINSREQ 225
 DB 154 CPVNVWEHQDSCYFWSHSGMSWAEAEKYCOLKNAHLVWVINSREQ 198
 RESULT 4
 LECH HUMAN
 ID LECH HUMAN STANDARD; PRT; 290 AA.
 AC P07306;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Asialoglycoprotein receptor 1 (Hepatic lectin H1) (ASGPR) (ASGPR-R).
 GN Name=ASGPR1;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=85130911; PubMed=2982798;
 RX Spiess M., Schwartz A.B., Lodish H.F.;
 RA "Sequence of human asialoglycoprotein receptor cDNA. An internal
 RT signal sequence for membrane insertion."
 RL J. Biol. Chem. 260:1979-1982 (1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86079574; PubMed=3753585; DOI=10.1016/0092-8674(86)90496-4;
 RA Spiess M., Lodish F.;
 RT "An internal signal sequence: the asialoglycoprotein receptor membrane
 anchor."
 RL Cell 44:177-185 (1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Wang H., Gao X., Li L., Lou H., Huang Y., Wang B., Han J.;
 RA "Human asialoglycoprotein receptor 1 gene is expressed in SH-SY5Y
 RT human neuroblastoma cells."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Udwin T.B., Tohiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [5]
 RP INTERACTION WITH LASS2.
 RX PubMed=11543633; DOI=10.1006/geno.2001.6614;
 RA Pan H., Qin W.-X., Huo K.-K., Wan D.-P., Yu Y., Xu Z.-G., Hu Q.-D.,
 RA Gu K.-T., Zhou X.-M., Jiang H.-Q., Zhang P.-P., Huang Y., Li Y.-Y.,
 RA Gu J.-R.;
 RT "Cloning, mapping, and characterization of a human homologue of the
 RT yeast longevity assurance gene LAG1."
 RL Genomics 77:58-64 (2001).
 CC -1- FUNCTION: Mediates the endocytosis of plasma glycoproteins to
 CC which the terminal sialic acid residue on their complex
 CC carbohydrate moieties has been removed. The receptor recognizes
 CC terminal galactose and N-acetylgalactosamine units. After ligand
 CC binding to the receptor, the resulting complex is internalized and
 CC transported to a sorting organelle, where receptor and ligand are
 CC disassociated. The receptor then returns to the cell membrane
 CC surface.
 CC -1- SUBUNIT: Interacts with LASS2.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed exclusively in hepatic parenchymal
 CC cells.
 CC -1- PTM: Phosphorylated on a cytoplasmic Ser residue.
 CC -1- MISCELLANEOUS: Calcium is required for ligand binding.
 CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on way
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: M10058; AA51785.1; -.
 DR EMBL: AB070933; BAB83508.1; -.
 DR EMBL: BC032130; AAH32130.1; -.
 DR PIR: A22509; LNHU1.
 DR PDB: 1DV8; X-ray; A=153-280.
 DR Genew: HGNC:742; ASGR1.
 DR H-InvDB: HIX0013487; -.
 DR MIM: 108360; -.
 DR GO: GO:0005887; C:integral to plasma membrane; TAS.
 DR GO: GO:0004873; F:asialoglycoprotein receptor activity; TAS.
 DR GO: GO:0006898; P:receptor mediated endocytosis; TAS.
 DR InterPro: IPR002353; AntifreezeZell.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C; 1.
 DR Pfam: PF03954; Lectin_N; 1.
 DR PRINTS: PR00356; ANTI-FREEZE2II.
 DR PROSITE: PS00615; C-TYPE-LECTIN_1; 1.
 DR PROSITE: PS00041; C-TYPE-LECTIN_2; 1.
 DR 3D-structure: Calcium; Endocytosis; Glycoprotein; Lectin;
 KW Phosphorylation; Receptor; Signal-anchor; Transmembrane.
 FT INIT_MET 0 0
 FT DOMAIN 1 39 Cytoplasmic (Probable).
 FT TRANSMEM 40 60 Signal-anchor for type II membrane
 FT protein (Probable).
 FT DOMAIN 61 290 Extracellular (Probable).
 FT DOMAIN 152 278 C-type lectin.
 FT SITE 4 7 Endocytosis signal (Potential).
 FT DISULFID 153 164 By similarity.
 FT DISULFID 181 276 By similarity.

FT	DISULFID	254	268	By similarity.
FT	CARBOHYD	80	80	O-linked.
FT	CARBOHYD	148	148	O-linked.
FT	TURN	155	156	
FT	STRAND	158	160	
FT	TURN	161	162	
FT	STRAND	163	167	
FT	STRAND	172	172	
FT	HELIX	174	183	
FT	TURN	184	185	
FT	STRAND	187	188	
FT	HELIX	194	204	
FT	STRAND	209	214	
FT	TURN	216	217	
FT	STRAND	221	222	
FT	TURN	223	224	
FT	STRAND	229	230	
FT	TURN	235	235	
FT	TURN	237	238	
FT	TURN	245	246	
FT	TURN	254	257	
FT	TURN	259	260	
FT	STRAND	263	266	
FT	TURN	268	269	
FT	STRAND	272	279	
SQ	SEQUENCE	290 AA; 33055 MW; B1897CE30DAE1586 CRC64;		
Query Match				
Best Local Similarity 52.8%; Score 918.5; DB 1; Length 290;				
Matches 171; Conservative 47; Mismatches 67; Indels 29; Gaps 5;				
QY	2	TRTYENFOYLENKVK-VQGFKNGLPQLQSLRLRSQCHLLSLGLGLLLVIVCVGF	60	
Db	1	TREYDQLHLDNEESDHQRLKRGPPPPQLQLRCSGPRLLLSLGLLLVIVCVIGS	60	
QY	61	QNSKFFORDLVLTDFSNFTSNTVARIQALTSQSSLEETIASLKAEEVGFQKQERQAGS	120	
Db	61	QNSQLEELRGLRETSNFTASTAQVKGSLTQGGNVGRKMSLSQLE--KQK-----	113	
QY	121	ELQEHHTQKAHLGHCPCFSVCVPVHSEMLLRVQQLVQDLKLTQVATLNNASTEGTC	180	
Db	114	DUSED-----HSSLLLVKQFVSDRLSLSCMAALQNGS-ERTC	152	
QY	181	CPVNWVEHODSCVYFSGHSGMSWAEKCYQLKNAHLVWINSREQNFQVQKYLGSAYTWNG	240	
Db	153	CPVNWVEHRSVCYFSGSKAWADADNYCRLEDAHLVVVTSWEEQKFVQHHIGPVNTWNG	212	
QY	241	LSDPEGAKWVDGTDYATQFQWKPQDDWQHGGLGGEDCAHFHPDGRWDDVCORPY	300	
Db	213	LHDQNGPWKVDGTDYETGFKWVRPQDDWYGHGLGGEDCAHFETDDGRWDDVCORPY	272	
QY	301	HWCEAGLGQTSQE	314	
Db	273	RWCETELDKASQE	286	
RESULT 5				
MMGL_RAT				
ID	MMGL_RAT	STANDARD;	PRT;	306 AA.
AC	P49301;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Macrophage asialoglycoprotein-binding protein (M-ASGP-Bp) (Macrophage			
DE	Galactose/N-acetylgalactosamine-specific lectin) (MWGL).			
GN	Name=Mg11; Synonyms=Mg1;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=90293078; PubMed=2358462;			

RA	Ii M., Kurata H., Itoh N., Yamashina I., Kawasaki T.;			
RT	"Molecular cloning and sequence analysis of cDNA encoding the			
RT	macrophage lectin specific for galactose and N-acetylgalactosamine.;"			
RL	J. Biol. Chem. 265:11295-11298(1990).			
RN	[2]			
RP	PRELIMINARY SEQUENCE OF 9-28.			
RX	MEDLINE=88339956; PubMed=3421964;			
RA	Ii M., Kawasaki T., Yamashina I.;			
RT	"Structural similarity between the macrophage lectin specific for			
RT	galactose/N-acetylgalactosamine and the hepatic asialoglycoprotein			
RT	binding protein.;"			
RL	Biochem. Biophys. Res. Commun. 155:720-725(1988).			
CC	-!- FUNCTION: Recognizes terminal galactose and N-acetylgalactosamine			
CC	units.			
CC	-!- SUBUNIT: Homooligomer.			
CC	-!- SUBCELLULAR LOCATION: Type II membrane protein.			
CC	-!- SIMILARITY: Contains 1 C-type lectin family domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement [see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch].			
CC	-----			
DR	EMBL; J05495; AAA41216.1; -.			
DR	PIR; A42230; A42230.			
DR	HSP; P07306; IDV8.			
DR	InterPro; IPR002353; AntifreezeII.			
DR	InterPro; IPR001304; Lectin C.			
DR	InterPro; IPR005640; lectin N.			
DR	InterPro; IPR009053; Prefoldin.			
DR	Pfam; PF00059; Lectin_C; 1.			
DR	Pfam; PF03954; Lectin_N; 1.			
DR	PRINTS; PR00356; ANTIFREEZEII.			
DR	SMART; SM00034; CLECT; 1.			
DR	PROSITE; PS00615; C-TYPE LECTIN 1; 1.			
DR	PROSITE; PS00041; C-TYPE LECTIN 2; 1.			
KW	Calcium; Direct protein sequencing; Glycoprotein; Lectin;			
KW	Signal-anchor; Transmembrane.			
FT	DOMAIN 1 37			
FT	TRANSMEM 38 58			
FT	DOMAIN 59 306			
FT	DOMAIN 174 300			
FT	DISULFID 175 186			
FT	DISULFID 203 298			
FT	DISULFID 276 290			
FT	CARBOHYD 76 76			
FT	CARBOHYD 168 168			
SQ	SEQUENCE 306 AA; 34242 MW; D68A5DFF0B9B8F13 CRC64;			
Query Match 51.0%; Score 887; DB 1; Length 306;				
Best Local Similarity 54.8%; Pred. NO. 7.4e-66;				
Matches 177; Conservative 38; Mismatches 80; Indels 28; Gaps 5;				
QY	1	MTRYENFOYLENKVKVQGFKNGLPQLQSLRLRSQCHLLSLGLGLLLVIVCVGF	60	
Db	1	MTMAYENFQNLGSEKQ--EAGKAPQSPFCNLSWTHLLLSGLSLLLVIVSIVGS	58	
QY	61	QNSKFFORDLVLTDFSNFTSNTVARIQALTSQSSLEETIASLKAEEVGFQKQERQAG--	118	
Db	59	QNSQLRRDLTLRTLTNTNTKAEQLALSRGDSLTGINSLKVEVDHGHQELQAGRG	118	
QY	119	-----VSELO-----EHTTQKAHLGHCPCFSVCVPVHSEMLLRVQQLVQDLKLTQCVAT	169	
Db	119	LSQKVASLESTVEKEQLRTDL-----SEITDRVQQLGDKLTKLTLCQLAS	164	
QY	170	LNNASTEGTCPPVNWVEHODSCVYFSGHSGMSWAEKCYQLKNAHLVWINSREQNFQV	229	
Db	165	LKNNGSAV-ACCPHLHMEHSGCYWFSQSGKWPEDKCYQLENSLVVNSLAEQNFQ	223	

```

QY 230 KYLGSAITWGLSDPEGAKWVDGTYATGQFQWKPQDDWQGHGLGGEDCAHFHPDG 289
Db 224 THMGSVVTWIGLTDQNGPRWVDGTYEKGFTTHWAPQPDWYGHGLGGEDCAHFSTDG 283
QY 290 RWNDDVCQRPYHWVCBAGLGQTS 312
Db 284 RWNDDVCQRPYRWCEMKLAKDS 306

RESULT 6
QY184
ID QY184 PRELIMINARY; PRT; 284 AA.
AC QY184;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Asialoglycoprotein receptor major subunit (Asialoglycoprotein receptor
DE 1).
GN Name=Asgr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ; TISSUE=Liver;
RX MEDLINE=20137499; PubMed=10675034; DOI=10.1016/S0378-1119(99)00493-X;
RA Soukharev S., Berlin W., Hanover J.A., Bethke B., Sauer B.;
RT "Organization of the mouse ASGR1 gene encoding the major subunit of
RT the hepatic asialoglycoprotein receptor.";
RL Gene 241:233-240(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Spatleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Tshiyuki S., Carninci P., Prange C.,
RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF182811; AAF29495.1; -
DR EMBL; BC022106; AAH22106.1; -
DR PIR; S29855; S29855.
DR HSSP; P07306; IDV8.
DR MGD; MGI-88081; Asgr1.
DR GO; GO:0016021; C: integral to membrane; TAS.
DR Pfam; PF00059; Lectin_C; 1.
DR PRINTS; PF03954; Lectin_N; 1.
DR SMART; PR00356; ANTIREEZEII.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PSS0041; C_TYPE_LLECTIN_2; 1.
KW Receptor.

```

```

SQ SEQUENCE 284 AA; 32591 MW; 671F043968047DB7 CRC64;
Query Match 50.5%; Score 877; DB 2; Length 284;
Best Local Similarity 51.3%; Pred. No. 4.6e-65;
Matches 160; Conservative 54; Mismatches 70; Indels 28; Gaps 4;

QY 1 MTRTYENFOYLENKVKVQGFQKNGPLQSLQRLQRLSGPCCHLLSLGLGLLLVIVCVGF 60
Db 1 MYKDYQDFHLDNDNDHQLRGPPPTPRLLQLRCLSGRLLLSLSLLVLLVVCVITS 60
QY 61 QNSKFORDLVTLTDFSNFTSNTVAEIQALTSQSSLEETIASLKAEVEGFQKQAGVS 120
Db 61 QNSQLREDLLALRQNFESNLTVSTEDQVKALSTQSSVGRKMKLVESKLE--KQK----- 113
QY 121 ELQEHRTQXAHGLGCHPCFSVCVPVHSEMLLRVQVLQDLKLTQCVATLNNASTEGTC 180
Db 114 DLTED-----HSLLLHVQLVSDVRSLSQMAAFRNGS--ERTC 152
QY 181 CPWNVVEHODSCYFWSHSGMSWAEAKYCOLKNAHLVWINSREQNFVOKYLGSAVTWG 240
Db 153 CPINWVEYEGSCYFWSVSRPWTEDKYCOLENAHLVVTSTRDEQNFQRHMGPLNTWIG 212
QY 241 LSDPEGAWKWVDGTYATGQFQWKPQDDWQGHGLGGEDCAHFHPDGRWDDVCQRPY 300
Db 213 LTDQNGPKWVDGTYETGTFQWNRPEQPDWYGHGLGGEDCAHFTTDCGRWDDVCRRPY 272
QY 301 HWVCEAGLGQTS 312
Db 273 RWCETKLDKAN 284

RESULT 7
LECH RAT STANDARD; PRT; 283 AA.
ID AC P02706;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Asialoglycoprotein receptor 1 (Hepatic lectin 1) (RHL-1) (ASGP-R)
DE (ASGPR).
GN Name=Asgr1; Synonyms=Asgr-1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86008335; PubMed=2995379;
RA Leung J.O., Holland E.C., Drickamer K.;
RT "Characterization of the gene encoding the major rat liver
RT asialoglycoprotein receptor.";
RL J. Biol. Chem. 260:12523-12527(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85063786; PubMed=6095287;
RA Holland E.C., Leung J.O., Drickamer K.;
RT "Rat liver asialoglycoprotein receptor lacks a cleavable NH2-terminal
RT signal sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:7338-7342(1984).
RN [3]
RP SEQUENCE OF 11-283 FROM N.A.
RX MEDLINE=87026895; PubMed=2945599;
RA Watts C.;
RT "Isolation and expression of cDNA clones for a rat liver
RT asialoglycoprotein receptor.";
RL Biosci. Rep. 6:527-534(1986).
CC -1- FUNCTION: Mediates the endocytosis of plasma glycoproteins to
CC which the terminal sialic acid residue on their complex
CC carbohydrate moieties has been removed. The receptor recognizes
CC terminal galactose and N-acetylgalactosamine units. After ligand
CC binding to the receptor, the resulting complex is internalized and
CC transported to a sorting organelle, where receptor and ligand are
CC disassociated. The receptor then returns to the cell membrane

```


Db 297 GRWNDDVCQRHYHWHICTELKAS 320

RESULT 9

LECH MOUSE

AC LECH MOUSE STANDARD; PRT; 283 AA.

DT 01-FEB-1994 (Rel. 28, Created)

DT 15-FEB-1998 (Rel. 36, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Asialoglycoprotein receptor 1 (Hepatic lectin 1) (MHL-1) (ASGP-R) (ASGP)

GN Name=Asgr1; Synonyms=Asgr-1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93176818; PubMed=8439566; DOI=10.1016/0167-4781(93)90300-3;

RA Takekawa R., Shinzawa K., Watanabe Y., Akaike T.;

RT "Determination of mouse major asialoglycoprotein receptor cDNA sequence.";

RL Biochim. Biophys. Acta 1172:220-222(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Liver;

RX MEDLINE=95047431; PubMed=7958950; DOI=10.1016/0378-1119(94)90694-7;

RA Monroe R.S., Huber B.E.;

RT "The major form of the murine asialoglycoprotein receptor: cDNA sequence and expression in liver, testis and epididymis.";

RL Gene 148:237-244(1994).

CC -1- FUNCTION: Mediates the endocytosis of plasma glycoproteins to which the terminal sialic acid residue on their complex carbohydrate moieties has been removed. The receptor recognizes terminal galactose and N-acetylgalactosamine units. After ligand binding to the receptor, the resulting complex is internalized and transported to a sorting organelle, where receptor and ligand are dissociated. The receptor then returns to the cell membrane surface.

CC -1- SUBUNIT: Interacts with LASS2 (By similarity).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein.

CC -1- TISSUE SPECIFICITY: Expressed exclusively in hepatic parenchymal cells.

CC -1- MISCELLANEOUS: Calcium is required for ligand binding.

CC -1- SIMILARITY: Contains 1 C-type lectin family domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D13517; BAA02734.1; -;

DR EMBL; U09362; AAB60441.1; -;

DR EMBL; U08372; AAB60440.1; -;

DR HGSP; P07306; 1DV8.

DR MGD; MGI:88081; Asgr1.

DR InterPro; IPR002353; AntifreezeZell.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR005640; Lectin_N.

DR Pfam; PF00059; Lectin_C; 1.

DR Pfam; PF03954; Lectin_N; 1.

DR PRINTS; SM00356; ANTIFREEZEII.

DR SMART; SM0034; CLECT; 1.

DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.

DR PROSITE; PS00441; C-TYPE LECTIN 2; 1.

KW Calcium; Endocytosis; Glycoprotein; Lectin; Phosphorylation; Receptor;

KW Signal-anchor; Transmembrane.

FT INIT_MET 0 0 By similarity.

FT DOMAIN 1 38 Cytoplasmic (Potential).

FT

FT TRANSMEM 39 59 Signal-anchor for type II membrane protein (Potential).

FT DOMAIN 60 283 Extracellular (Potential).

FT SITE 4 7 C-type lectin.

FT DISULFID 152 163 Endocytosis signal (Potential).

FT DISULFID 180 275 By similarity.

FT DISULFID 253 267 By similarity.

FT CARBOHYD 74 74 N-linked (GLNAC. . .) (Potential).

FT CARBOHYD 77 77 N-linked (GLNAC. . .) (Potential).

FT CARBOHYD 145 145 N-linked (GLNAC. . .) (Potential).

FT CONFLICT 150 150 I -> T (in Ref. 1).

SQ SEQUENCE 283 AA; 32472 MW; 982A5D305AAE0D8F CRC64;

Query Match 49.8%; Score 866; DB 1; Length 283;

Best Local Similarity 50.8%; Pred. No. 3.8e-64;

Matches 158; Conservative 54; Mismatches 71; Indels 28; Gaps 4;

QY 2 TRTYENFOYLENKVKVQGFKNGLPLQSLQRLQSRGPGCHLLSLGLGLLLVIVCVVGFQ 61

Db 1 TKDYQFQHLNDNDHQLRRGPPPTRLQLRCGSRLLSSLSILLVVCVITSQ 60

QY 62 NSKFORDLVTRTDFSNFTSNTVAEIQALTSQGSLEETIASLKAEEVGFQERQAGVSE 121

Db 61 NSQLREDLLALRQNFSLTVSTEDQVKALSTQGSVGRKMKLVESKLE--KQK-----D 113

QY 122 LQHTTQKAHLGHCHPCSPVCVPHSEMILLRVQVODLKKLTCQVATLNNASTETGCC 181

Db 114 LTED-----HSSLHLRVQLVSDVRSLSQMAAFRNGS-ERICC 152

QY 182 PVNWEHQDSCYWFSGHSMWAEAEKYCOLKNAHLVVINSREQNFQVKYLGSAYTWMLG 241

Db 153 PINWVEYEGSCYWFSSSVRPMTADKYQLENLAHLVVVTSRDBQNFQRMGPTNWTIGL 212

QY 242 SDPEGAWKWDGTDYATGFGQWKPQDQDDQHGGLGGEDCAHFHPDGRWDDVCCORPYH 301

Db 213 TDQNGPKWVDGTDYETGFGQWPEQPDNWKYHGLGGEDCAHFTTDRWNDDVCRPYR 272

QY 302 WVCEAGLGQTS 312

Db 273 WVCETKLDKAN 283

RESULT 10

Q91YT3

ID Q91YT3 PRELIMINARY; PRT; 304 AA.

AC Q91YT3;

DT 01-DEC-2001 (T-EMBLrel. 19, Created)

DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)

DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

DE Macrophage galactose N-acetyl-galactosamine specific lectin 1.

GN Name=Mgl1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N; TISSUE=Mammary tumor;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,

RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,


```

QY 179 TCCPVNVVEHODSCYWFSGMSWAEKCYCOLKNAHLVINSREBQNFQKYLGSAYTW 238
Db 171 ACCPLHTEHEGSCYWFSEKSWPADKYCRLENSHLVVNSLEBQNFQNLANVSW 230
QY 239 MGLSDPEGAKWVDGTDYATGFQWPKPGQDDQGHGLGGEDCAHFPDPGRWDDVCQR 298
Db 231 IGLTDQNGPWRWDGTDGDFEKGFFKNAPLQPDNFWGHGLGGEDCAHITGGPWNDDVCQR 290
QY 299 PYHWVCEAGLGQTS 312
Db 291 TFRWICEMKLAKES 304

RESULT 12
Q8BUD5 PRELIMINARY; PRT; 361 AA.
AC Q8BUD5;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Mus musculus 10 days lactation, adult female mammary gland cDNA, RIKEN
DE full-length enriched library, clone:D730047H02 product:similar to
DE MACROPHAGE GALACTOSE N-ACETYL-GALACTOSAMINE SPECIFIC LECTIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=99279233; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Akiyama J., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [5]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).

```

```

RN [6]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai T.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sugabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK085751; BAC39530.1; -.
DR HSSP; P07306; 1DV8.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; LECTIN_C.
DR InterPro; IPR005640; LECTIN_N.
DR InterPro; IPR009053; Prefoldin.
DR Pfam; PF00059; LECTIN_C; 1.
DR Pfam; PF03954; LECTIN_N; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS0041; C_TYPE_LECTIN_2; 1.
KW LECTIN.
SQ SEQUENCE 361 AA; 41010 MW; AACBBY748C66A231 CRC64;

Query Match 45.5%; Score 791; DB 2; Length 361;
Best Local Similarity 47.0%; Pred. No. 9.8e-58;
Matches 158; Conservative 40; Mismatches 86; Indels 52; Gaps 3;

QY 25 LPLQLQLRSLRQPCCHLLSLGLGLLLVLCVVGQFQNSKFORDLVLRTRDFSNFTSNTV 84
Db 18 IPSQSFELWILSWTHLLFSLGLSLLLVVISVIGSQNSQLRDLGLTRAILDNTSKIK 77
QY 85 AEIQALTSQGSLEETIASLKAEVGFQKRRQAGVSELOHTTKAHLGHCHPCPVCVP 144
Db 78 AEFQSLDSRADNFEKGISSIKVDVEDHRLQBLAQGRDLSQKVTLSLESTLEKREKALKTDL- 136
QY 145 VHSEMLLRVQQLVODLKLKLTQVATLNNASTGTCTCPVNVVEHODSCYWFSGMSWAE 204
Db 137 --SDLTDHVVQQLTDLKALTQCLANLKNNGS-EVACCPHLHTEHEGSCYWFSEKSWPE 193
QY 205 AEKYCOLKNAHLVINSREBQNFQKYLGSAYTWMLGSDPEGAKWVDGTDYATG----- 259
Db 194 ADKYCELENSHLVVNSLEBQNFQNLANVLSWMGLTDQNGPWRWDGTDGDKGPKYVC 253
QY 260 -----FQNKPGQDDQWQGHGL 276
Db 254 RLQLAPLYLGLSYLFSIFSDPRDLGPGSGNMADGQIWSAQFFIFRNWRPLQPDNWHGML 313
QY 277 GGEDCAHFPDPGRWDDVCQRYHWVCEAGLGQTS 312
Db 314 GGGEDCAHFSYDGRWDDVCQRYHWVCEAGLGQTS 349

RESULT 13
Q7Z4G9 PRELIMINARY; PRT; 306 AA.
AC Q7Z4G9;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE HbXAg-binding protein.
GN Name=HbXBP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```


FT	CARBOHYD	97	97	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	165	165	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	298	298	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	301 AA;	34907 MW;	3A29F1AFBA68F298 CRC64;
Query Match 38.6%; Score 671.5; DB 1; Length 301;				
Best Local Similarity 46.1%; Pred. No. 7.8e-48;				
Matches 130; Conservative 42; Mismatches 75; Indels 35; Gaps 4;				
Qy	24	PLPLQSLQLRLRSGPC	HLLSLGLGLLLVITCVVGFQNSKFQ	RDILVTLRTDFSNFTSNT 83
Db	48	PPP-----QRLCST	FRLSLLALAFNILLVVICVSSQSIQ	OEFRILKETFSNFSST 102
Qy	84	VAEIQALTSQSSLEET	IASLKAIEVGFQERQAGVSELQEH	TTQKAHLGHCPHCPSVCV 143
Db	103	LMEFGALDTLGGST	NAILTSWLAQLEEKQQLKA-----	DHSTLLFLHKFP----- 149
Qy	144	PVHSEMLLRVQQLV	QDLKKLTQVATLNNNA	STEGTCCPNNVVEHQDSCYWFSGMSWA 203
Db	150	-----MDLRTL	TCQLAYFQSGNT---ECCP	VNWVEFGGSCYWFSGDGLTWA 192
Qy	204	EAEKCYQLKNAHLV	VINSREEQNFVOKYLGSAVTW	GLSDPEGAMKWVDGTDYATGFQNW 263
Db	193	EADQYCOLENAHLV	INSREEQDFVVKHRSQPHIW	IGLTDGSMKWVDGTDYRSNRYNW 252
Qy	264	KPGQPDQWQGHGLG	GGEDCAHFHPDGRW	NDVQCQRYHWVCE 305
Db	253	AFTQPDNMQGHEQ	GGEDCAEILSDGHWN	DNFCQVNRWVCE 294

Search completed: May 27, 2005, 12:48:52
Job time : 68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 12:44:45 ; Search time 17 Seconds
(without alignments)
1788.500 Million cell updates/sec

Title: US-10-829-107-4
Perfect score: 1738
Sequence: 1 MTRTYENFOYLENKVKVQGF.....QRPYHWCEAGLGQTSQESH 316

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 79: *
1: Pirl: *
2: Pirl2: *
3: Pirl3: *
4: Pirl4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	923.5	53.1	291	1 LNHU1	hepatic lectin H1
2	887	51.0	306	2 A42230	lectin M-ASGP-BP p
3	879	50.6	284	1 LNR7L	hepatic lectin - r
4	877	50.5	284	2 S29855	asialoglycoprotein
5	804	46.3	304	2 JX0209	lectin, galactose/
6	754.5	43.4	311	1 LNHU2A	asialoglycoprotein
7	671.5	38.6	301	2 S13165	asialoglycoprotein
8	655.5	37.7	301	1 LNR72	hepatic lectin 2 -
9	326.5	18.8	550	2 A28166	Kupffer cell recep
10	316	18.2	404	2 A46274	HIV gp120-binding
11	313	18.0	742	2 JC7595	scavenger receptor
12	288.5	16.6	309	1 S34198	IGF Rc receptor II
13	287	16.5	331	1 LNM5ER	IGF Rc receptor, I
14	281	16.2	207	1 LNC4L	hepatic lectin - c
15	274.5	15.8	321	1 LNHUER	IGF Rc receptor II
16	259	14.9	155	2 S78774	perleucan - Halloti
17	253	14.6	237	2 JC7608	type II lectin-like
18	243.5	14.0	2397	1 A55535	versican precursor
19	240	13.8	1340	2 A39808	proteoglycan core
20	240	13.8	2327	2 T42630	aggreccan - bovine
21	239.5	13.8	2409	1 A60979	versican precursor
22	239.5	13.8	3562	2 A47171	chondroitin sulfat
23	239	13.8	1643	2 T14274	versican precursor
24	238.5	13.7	1268	2 S52781	neurocan - mouse
25	237.5	13.7	3381	2 T42389	versican precursor
26	236.5	13.6	2109	1 T50421	aggreccan precursor
27	232	13.3	2124	2 A28452	proteoglycan core
28	228	13.1	2152	1 A55182	aggreccan precursor
29	228	13.1	2415	1 A39086	aggreccan precursor

30	226.5	13.0	1257	2 S28764	neurocan precursor
31	223.5	12.9	912	2 A54423	brevican precursor
32	222.5	12.8	330	2 T46256	brevican - human (
33	219.5	12.6	883	2 S57653	brevican precursor
34	218.5	12.6	883	2 S49126	brevican precursor
35	204	11.7	162	1 LNR01	lectin BRA3-1 prec
36	198.5	11.4	147	2 A26697	echinoidin - sea u
37	196	11.3	162	1 LNR03	lectin BRA3-2 prec
38	188	10.8	551	2 I46709	endothelial leukoc
39	187.5	10.8	257	2 I50146	gene 17.5 protein
40	186	10.7	199	2 JH0822	lymphocyte early a
41	185	10.6	172	2 S32489	lectin - Iberian r
42	181	10.4	163	1 A34313	antifreeze protein
43	180	10.4	372	2 JC5377	L-selectin precurs
44	179	10.3	262	2 T34115	hypothetical prote
45	178	10.2	612	2 B42755	E-selectin precurs

ALIGNMENTS

RESULT 1

LNHU1

hepatic lectin H1 - human

N;Alternate names: asialoglycoprotein receptor H1 (ASGP-H1)

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004

C;Accession: A22509

R;Spiess, M.; Schwartz, A.L.; Lodish, H.F.

J. Biol. Chem. 260, 1979-1982, 1985

A;Title: Sequence of human asialoglycoprotein receptor cDNA: an internal signal sequence

A;Reference number: A22509; MUID:85130911; PMID:2382798

A;Accession: A22509

A;Molecule type: mRNA

A;Residues: 1-291 <SPI>

A;Cross-references: UNIPROT:P07306; GB:M10058; NID:G179078; PID:AAA1785.1; PID:G179079

C;Comment: This receptor is expressed in mammals exclusively in hepatic parenchymal cells

C;Comment: By homology with the H1 receptor, the initiator Met is removed after translat

C;Comment: A cytoplasmic serine residue is phosphorylated.

C;Genetics:

A;Gene: GDB:ASGR1

A;Cross-references: GDB:118754; OMIM:108360

A;Map position: 17p13-17p11

C;Superfamily: hepatic lectin; C-type lectin homology

C;Keywords: endocytosis; glycoprotein; lectin; phosphoprotein; receptor; transmembrane p

F;2-291/Product: hepatic lectin H1 #status predicted <MAT>

F;2-40/Domain: intracellular #status predicted <INT>

F;41-59/Domain: transmembrane #status predicted <TM>

F;60-231/Domain: extracellular #status predicted <EXT>

F;154-277/Domain: C-type lectin homology <LCH>

F;79,147/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.1%; Score 923.5; DB 1; Length 291;

Best Local Similarity 54.6%; Pred. No. 1.5e-67;

Matches 172; Conservative 47; Mismatches 67; Indels 29; Gaps 5;

Qy 1 MTRTYENFOYLENKVK-VQGFKNGLPLQSLQLRSGPCHLLSLGLGLLLVLCWG 59

Db 1 MTKYQDLQHLNDESDHQLRKGPPOPQLQLCSGPRLLLSLSLLLVVVCVIG 60

Qy 60 FQNSKQFDLVLTDTFSNFTNTVAEIQALTSQSSLEETIASLKAEVSGFKQERQAGV 119

Db 61 SONSQLELRGLURETFNSNFTASTAQVGLSTQGGVGRKMKLSLSQLE--KQOK---- 114

Qy 120 SELQHTTQKAHLGHCPSPVCVPVHSEMLLRVQQLVQDLKKLTCCQVATLNNASTEGT 179

Db 115 -DUSED-----HSSLLLVHVKQFVSDRLSLSCQMAALQNGS-ERT 152

Qy 180 CCPVNWVEHQDSCYWFSGHSGMWAEEKYCQLKNAHLVWINSREEQNFVKYLGSAVTWM 239

Db 153 CCPVNWVEHRSYWFSGKAWADADNYCRLEDAHLVVVTWSWEEKFVQHHGPNVTWM 212

Qy 240 GLSDPGEAKWKVDGTATYATGFQNKPKQPDQWQHGGLGGEDCAHFHPDGRWDDVCQRP 299

Db 213 GLHDQNGPFWKWDGTDYETGKWRPEQPDWYGHGLGGGDCAHFTDDGRWDDVQRP 272
QY 300 YHWVCEAGLGQTSQE 314
Db 273 YRWVCETELDKASQE 287

RESULT 2
A42230
lectin M-ASGP-BP precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 09-Jul-2004
C:Accession: A42230
R;Ii, M.; Kurata, H.; Itoh, N.; Yamashina, I.; Kawasaki, T.
J. Biol. Chem. 265, 11295-11298, 1990
A:Title: Molecular cloning and sequence analysis of cDNA encoding the macrophage lectin
A:Reference number: A42230; MUID:90293078; PMID:2358462
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-306 <L1A>
A:Cross-references: UNIPROT:P49301; GB:J05495; NID:g204302; PIDN:AAA41216.1; PID:g204303
C:Superfamily: hepatic lectin; C-type lectin homology
F:175-298/Domain: C-type lectin homology <LCH>

Query Match 51.0%; Score 887; DB 2; Length 306;
Best Local Similarity 54.8%; Pred. No. 1.5e-64;
Matches 177; Conservative 38; Mismatches 80; Indels 28; Gaps 5;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPQLQSLQRLQRSGFCHLLLSGLGLGLLLVLCVVG 60
Db 1 MTWAYENFONLGSBEKQ--EAGKAPQSFCLNLTLSWTHLLFSGLGLLLVVISVIGS 58
QY 61 QNSKFQDVLTRTDFNSFTNTVAETQALTSQSSLEETIASLKAEVGEFKQBRQAG-- 118
Db 59 QNSQLRDLLETTRTTLDTNTTNTKAELQALASRGDSLQGTGINSLKVEVDHGOELQAGRG 118
QY 119 ----VSLQ-----EHTTQKAHLGHCPHCPSCVCPVHSEMLLRVQQLVQDLKKLTCCOVAT 169
Db 119 LSQKVASLESTVEKQETRLTDL-----SEITDRVQQLGKDLKTLTCLQAS 164
QY 170 LNNNASTEGTCCPVNWEYHODSCYWFHSGMSWAEAEKYCOLKNAHLVWVINSREEQNFPVQ 229
Db 165 LKNGSAV-ACCP LHWMEHSGCYWFSGKGPWEADKYQLENSNLVVNSLAEQNFLQ 223
QY 230 KYLGSAVTWGLSDPEGAWKVDGTDYATGFQNWKPQDPDDWQHGLGGGDCAHFHPDG 289
Db 224 THMSVVVTWGLTDQNGPWRVWDGTDYEKGFTHWAPKQPDNWNWYGHGLGGGDCAHFTSDG 283
QY 290 RWNDVDCQRPYHWVCEAGLGQTS 312
Db 284 RWNDVDCQRPYRWVCEMKLAKDS 306

RESULT 3
LNRTL
hepatic lectin - rat
N:Alternate names: ASGP; asialoglycoprotein receptor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004
C:Accession: A92497; A94020; B94020; A54727; A03166
R;Leung, J.O.; Holland, E.C.; Drickamer, K.
J. Biol. Chem. 260, 12523-12527, 1985
A:Title: Characterization of the gene encoding the major rat liver asialoglycoprotein re
A:Reference number: A92497; MUID:86008335; PMID:2995379
A:Accession: A92497
A:Molecule type: DNA
A:Residues: 1-284 <LEU>
A:Cross-references: UNIPROT:P02706; GB:K02817; NID:g2066646; PIDN:AAA42037.1; PID:g206647
R;Holland, E.C.; Leung, J.O.; Drickamer, K.
Proc. Natl. Acad. Sci. U.S.A. 81, 7338-7342, 1984
A:Title: Rat liver asialoglycoprotein receptor lacks a cleavable NH-2-terminal signal se

Db 61 LRDLGLTRATLDNTTSKIKAEFQSLDRADSFKEGISLKVVDVHEDHQELQAGRDLQK 120

Qy 119 VSELOEHTTKAHILGHCHPCPSVCVPVHSEMLLRVQQLVQDILKLTCQVATLNNASTEG 178
| : | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
121 VTSL-ESTVEKRE-----QALKTLDSLTDTHVQQLRKDKLKALTCQLANLNKNGS-EV 170

Qy 179 TCCPVNVVEHQDSYCFWRSHSGMSWAAEAKEYCOLKNAHLVVINSREONFVKYLGSAYTW 238
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
171 ACCPLHWTEHGSCYWFSESSEKSWEADKYCRLENSHLVLVVNLSBEQNFLORLANVSW 230
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
239 MGLSDPEGAKWKVDGTGYATGFQNWKKPQDPDWQOQHGLGGGBDCAHFHFDGRWNDDVCOR 298
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
231 IGLTDQNGPRWVDUGTDFEKGFKWAFLPDPNFWFGHGLGGGEDCAHIITCGPWNDVDCQR 290
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 299 PYHWVCEAGLGQTS 312
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
291 TFRWICEMKLAKES 304
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 6
LNHUZA
asialoglycoprotein receptor H2a - human
N/Alternate names: hepatic lectin H2a
N/Contents: ASGPR; asialoglycoprotein receptor H2a; asialoglycoprotein receptor H2b
C/Species: Homo sapiens (man)
C/Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 09-Jul-2004
C/Acession: A25179; A39100; B39100; I37995; A49466; B49466; S14525
R/Spies, M.; Lodish, H.F.
Proc. Natl. Acad. Sci. U.S.A. 82, 6465-6469, 1985
A/Title: Sequence of a second human asialoglycoprotein receptor: conservation of two rec
A/Reference number: A25179; MUID:86016723; PMID:3863106
A/Acession: A25179
A/Molecule type: mRNA
A/Residues: 1-311 <SPI>
A/Cross-references: UNIPROT:P07307; GB:M11025; NID:g179080; PIDN:AAB59519.1; PID:g179081
R/Lederkremer, G.Z.; Lodish, H.F.
J. Biol. Chem. 266, 1237-1244, 1991
A/Title: An alternatively spliced minixon alters the subcellular fate of the human asia
A/Reference number: A39100; MUID:91093236; PMID:1985943
A/Acession: A39100
A/Molecule type: DNA; mRNA
A/Residues: 69-99 <LED>
A/Cross-references: GB:M38420; NID:g184395
A/Acession: B39100
A/Molecule type: DNA; mRNA
A/Residues: 69-81,87-99 <LE2>
A/Cross-references: GB:M38420; NID:g184395
R/Paletta, E.; Stockert, R.J.; Racevsakis, J.
Hepatology 15, 395-402, 1992
A/Title: Differences in the abundance of variably spliced transcripts for the second asi
A/Reference number: I37995; MUID:92184202; PMID:I371982
A/Acession: I37995
A/Status: preliminary; translated from GB/EMBL/DBBJ
A/Molecule type: mRNA
A/Residues: 1-23,43-81,87-311 <PAI>
A/Cross-references: EMBL:X55283; NID:G34354; PIDN:CRAA38997.1; PID:g34355
R/Yuk, M.H.; Lodish, H.F.
J. Cell Biol. 123, 1735-1749, 1993
A/Title: Two pathways for the degradation of the H2 subunit of the asialoglycoprotein rec
A/Reference number: A49466; MUID:94103329; PMID:8276894
A/Acession: A49466
A/Molecule type: protein
A/Residues: 78-98 <YUK>
A/Acession: B49466
A/Molecule type: protein
A/Residues: 87-98 <YU2>
C/Comment: The functioning ligand-binding unit of this receptor is thought to be at least
C/Genetics:
A/Gene: GDB:ASGR2; L-H2
A/Cross-references: GDB:118755; OMTM:108361
A/Map position: 17p13-17p11
C/Superfamily: hepatic lectin; C-type lectin homology

C:Keywords: alternative splicing; endocytosis; glycoprotein; lectin; liver; phosphoprotein
F:1-311/Product: asialoglycoprotein receptor H2a #status predicted <MAT1>
F:1-81,87-311/Product: asialoglycoprotein receptor H2b #status predicted <MAT2>
F:1-58/Domain: intracellular #status predicted <INT>
F:1-23,43-81,87-311/Product: asialoglycoprotein receptor H2c #status predicted <MAT3>
F:59-78/Domain: transmembrane #status predicted <TM>
F:79-311/Domain: extracellular #status predicted <EXT>
F:177-300/Domain: C-type lectin homology <LCH>
F:102,170,305/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.4%; Score 754.5; DB 1; Length 311;
Best Local Similarity 47.8%; Pred. No. 8.9e-54;
Matches 143; Conservative 44; Mismatches 79; Indels 33; Gaps 4;

QY 20 FKNGPLPQLSLQRLSGPCHLLLSGLGLLLVIVCVGFQF-----SKFQDLVTLRT 74
DB 39 FLKGPAPQPLAQLRCSWCFSLALSFNILLVIVCVTGSSEGRGQLOAEIRSLKE 98
QY 75 DFSNFTSNTVAIEQALTSQSSLEETIASLKAIEVGFQKQERQAGVSEIQEHTTQKAHLGH 134
DB 99 AFSNFSSTLITEQAISTHGGVGDKITSLGAKLEKQOQDLKA-----DHDALLFHLKH 152
QY 135 CHPCFSCVCPVHSEMLLRVQVLVDLKLTCQVATLNNASTEGTCCPVNVVEHQDSCYW 194
DB 153 FP-----VDLRFVACQMELLHNSGS-QRTCCPVNVVEHQSCYW 190
QY 195 FSHSGMWAEAEKYCOLKNAHLVWVINSREQNFVKYLGSAVTWGLSDPEGAKWVDGT 254
DB 191 FSHSGKAWAEAEKYCOLENAHLVWVINSWEEQFIVQHTNPFTWIGLTDSDGSKWVDGT 250
QY 255 DYATGQFQNNKPGQDDWQHGGLGGEDCAHFPHDGRWDDVCQRPYHMYCEAGLQTSQ 313
DB 251 DYRHNKNAVTPQDNWHGHELGSGEDCEVEQPDGRWDDFCQLQYRVVCEKRRNATGE 309

RESULT 7
S13165
N:Alternate names: asialoglycoprotein receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S13165
R:Sanford, J.P.; Doyle, D.
Biochim. Biophys. Acta 1087, 259-261, 1990
A:Title: Mouse asialoglycoprotein receptor cDNA sequence: conservation of receptor genes
A:Reference number: S13165; MUID:91027942; PMID:2223888
A:Accession: S13165
A:Molecule type: mRNA
A:Residues: 1-301 <SAN>
A:Cross-references: UNIPROT:P24721; EMBL:X53042; NID:g53104; PIDN:CAA37211.1; PID:g531105
C:Superfamily: hepatic lectin; C-type lectin homology
C:Keywords: glycoprotein; liver; transmembrane protein
F:170-293/Domain: C-type lectin homology <LCH>

Query Match 38.6%; Score 671.5; DB 2; Length 301;
Best Local Similarity 46.1%; Pred. No. 4.8e-47;
Matches 130; Conservative 42; Mismatches 75; Indels 35; Gaps 4;

QY 24 PPLQSLQLRLSGPCHLLLSGLGLLLVIVCVGFQFQSKFQDLVTLRTDFSNFTSNT 83
DB 48 PFP-----QLCSTFRLSLALAFNILLVIVCVSSQSIQLQEFRTLTKEFTSNFSST 102
QY 84 VAEIQALTSQSSLEETIASLKAIEVGFQKQERQAGVSEIQEHTTQKAHLGHCPHCV 143
DB 103 LMEFGALDTLGGSTNAILTSWLAQLEEKQOQDLKA-----DHSLLFLHLKHPF----- 149
QY 144 PVHSEMLLRVQVLVDLKLTCQVATLNNASTEGTCCPVNVVEHQDSCYWFSSHGMSWA 203
DB 150 -----MDRLTITCQLAYFQNGT---ECCPVNVVEFGGSCYWFSDGLTWA 192
QY 204 EAEKYCOLKNAHLVWVINSREQNFVKYLGSAVTWGLSDPEGAKWVDGTATGATGQW 263
DB 193 EADQYCOLENAHLVWVINSREQDFVVKHRSQFHIWITGLTRDGSWKWVDGTGYRNRNW 252

QY 264 KPGQDDWQHGGLGGEDCAHFPHDGRWDDVCQRPYHMYCE 305
DB 253 AFTQPDNWNQHGEGGECBAEILSDGHWNDFCQOVNRWVCE 294

RESULT 8
LNRT2
N:Alternate names: asialoglycoprotein receptor RHL-2/3 (ASGP-R2/3)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 09-Apr-1998 #text_change 09-Jul-2004
C:Accession: B28462; A28462; A31601; A26888; A25417
R:Halberg, D.F.; Wager, R.E.; Farrell, D.C.; Hildreth IV, J.; Quesenberry, M.S.; Loeb, J.
J. Biol. Chem. 262, 9828-9838, 1987
A:Title: Major and minor forms of the rat liver asialoglycoprotein receptor are independent
A:Reference number: A28462; MUID:87250656; PMID:3597443
A:Accession: B28462
A:Molecule type: mRNA
A:Residues: 1-301 <HAL>
A:Cross-references: UNIPROT:P08290; GB:J02762; NID:g205162; PIDN:AAA41522.1; PID:g205163
A:Accession: A28462
A:Molecule type: protein
A:Residues: 88-96,'X',98-118,'X',120;129-158;177-182,'X',184,'X',186-189;192-290,'C',292
R:Sanford, J.P.; Elliott, R.W.; Doyle, D.
DNA 7, 721-728, 1988
A:Title: Asialoglycoprotein receptor genes are linked on chromosome 11 in the mouse.
A:Reference number: A31601; MUID:89170119; PMID:3234178
A:Accession: A31601
A:Molecule type: mRNA
A:Residues: 1-301 <SAN>
A:Cross-references: GB:X07636; NID:g57066; PIDN:CAA30476.1; PID:g57067
R:McPhaul, M.; Berg, P.
Mol. Cell. Biol. 7, 1841-1847, 1987
A:Title: Identification and characterization of cDNA clones encoding two homologous proteins
A:Reference number: A26888; MUID:87257885; PMID:3600647
A:Accession: A26888
A:Molecule type: mRNA
A:Residues: 1-152,'A',154-201,'I',203-259,'C',261-301 <MCP>
A:Cross-references: GB:M16347; NID:g206648; PIDN:AAA42038.1; PID:g206649
A:Note: the authors translated the codon GCA for residue 153 as Arg and ATT for residue 154 as Thr
R:Drickamer, K.; Mamon, J.F.; Binns, G.; Leung, J.O.
J. Biol. Chem. 259, 7770-7778, 1984
A:Title: Primary structure of the rat liver asialoglycoprotein receptor: structural evidence
A:Reference number: A25417; MUID:84111554; PMID:6319386
A:Accession: A25417
A:Molecule type: protein
A:Residues: 201-259,'C',261-281,'ND',284-301 <DRI>
C:Comment: Calcium is required for ligand binding.
C:Superfamily: hepatic lectin; C-type lectin homology
C:Keywords: endocytosis; glycoprotein; lectin; liver; receptor; transmembrane protein
F:2-60/Domain: intracellular #status predicted <INT>
F:61-77/Domain: transmembrane #status predicted <TRN>
F:78-301/Domain: extracellular #status predicted <EXT>
F:170-293/Domain: C-type lectin homology <LCH>
F:97,119,165/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.7%; Score 655.5; DB 1; Length 301;
Best Local Similarity 40.1%; Pred. No. 9.5e-46;
Matches 130; Conservative 55; Mismatches 90; Indels 49; Gaps 5;

QY 1 MTRTVENFQYLENVK-----VQGFK-----NGPLPQLSLQRLSGPCHL 41
DB 1 MEKQFDIQLQDSENDHQLIGDEQSGSHVQNLATENPRWCGQPPSRFPORLCSKFLS 60
QY 42 LLSLGLGLLLVIVCVGFQFQSKFQDLVTLRTDFSNFTSNTVAIEQALTSQSSLEETI 101
DB 61 LLALAFNILLVIVCVSSQSQMLQKEFTLTKEFTSLNFTSLTLMFEKALDSHGGSNDNL 120
QY 102 ASLKAIEVGFQKQERQAGVSEIQEHTTQKAHLGHCPHCVCPVHSEMLLRVQVLVDL 161
DB 121 TSWETILEKKORDIKA-----DHSLLFLHLKHPF-----LDLR 153


```
QY 162 KLTQVATLNNASTGTCPCPVNVEHQDSYFWSHSGMSWAEAEKYCOLKNAHLVINS 221
Db 154 TLTCQALFFLSNGT---ECCPVNVEFGGSCYFSDRLGTWAEADQYQCMENAHLLVINS 210
QY 222 REBQNFQKYLGSAYTWMGLSDPEGAWKWDGTDYATGFQNFQWKPQDQDQWQHGGLGGED 281
Db 211 REBQEFVVRGAFHIGLTKDGSWKWDGTEYRSNFKNWAFTQPDNQQHGGESD 270
QY 282 CAHFHPDGRWDDVCQRPYHWCE 305
Db 271 CABILSDGLWDDNFCQOVNRWACE 294

RESULT 9
A28166
Kupffer cell receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C;Accession: A38674; A28166
R;Hoyle, G.W.; Hill, R.L.
J. Biol. Chem. 286, 1850-1857, 1991
A;Title: Structure of the gene for a carbohydrate-binding receptor unique to rat Kupffer
A;Reference number: A38674; MUID:91107689; PMID:1846367
A;Accession: A38674
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-550 <HO2>
A;Cross-references: UNIPROT:P10716; GB:M55532; NID:g203362; PIDN:AAA40892.1; PID:g203363
R;Hoyle, G.W.; Hill, R.L.
J. Biol. Chem. 263, 7487-7492, 1988
A;Title: Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor
A;Reference number: A28166; MUID:88227939; PMID:2836387
A;Accession: A28166
A;Molecule type: mRNA
A;Residues: 1-550 <HOY>
A;Cross-references: GB:J03734; NID:g205050; PIDN:AAA41472.1; PID:g205051
A;Keywords: transmembrane protein
C;Keywords: Domain: C-type lectin homology <LCH>
F;412-536/Domain: C-type lectin homology <LCH>

Query Match 18.8%; Score 326.5; DB 2; Length 550;
Best Local Similarity 31.1%; Pred. No. 1.1e-18;
Matches 79; Conservative 32; Mismatches 94; Indels 49; Gaps 8;

QY 62 NSKFO---RDVLTDFSNFTNTVAEIOALTSQSSLEETIASLKAEVEGKQERQAG 118
Db 323 NGKLDSSRELQTLRDLSD-----DVSALKSNVQMLQSLQKAKAEVQSLKGLEA- 373
QY 119 VSELQHTTQKAHLGHCPCPSVCVPVHSEMLRVQQLVQDLKLTQCVATLNNASTEG 178
Db 374 -----TKTLAAKIQQQSDLEALQKAVAAHTQGOQTQN 406
QY 179 TCCPV---NWVEHQDSYFWSHSGMSWAEAEKYCOLKNAHLVINSREEQNFVQKYLGSA 235
Db 407 QVQLQLIMQDWKYNKGFYFSDKSKWHEAENFCVSGAHLASVTSQEEQAFVLQITNAV 466
QY 236 YTMWGLSD--PEGAWKWDGT--DYATGFQNFQWKPQDQDQWQHGGLGGEDCAHFHPDGRW 291
Db 467 DHWIGLTDQCTEGNWRWVDGTPEDYVQSRFRFKGQPDNWR-HGNGEREDCVHL--QRMW 523
QY 292 NDDVCQRPYHWCE 305
Db 524 NDMACGTAYNWNCK 537

RESULT 10
A46274
HIV gp120-binding C-type lectin - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46274
R;Curtis, B.M.; Scharnowske, S.; Watson, A.J.
Proc. Natl. Acad. Sci. U.S.A. 89, 8356-8360, 1992
A;Title: Sequence and expression of a membrane-associated C-type lectin that exhibits CD
```

```
A;Reference number: A46274; MUID:92390446; PMID:1518869
A;Accession: A46274
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-404 <CUR>
A;Cross-references: UNIPROT:Q9NNX6
A;Experimental source: placenta
A;Note: sequence extracted from NCBI backbone (NCBIN:113134, NCBIIP:113135)
F;256-377/Domain: C-type lectin homology <LCH>
```

```
Query Match 18.2%; Score 316; DB 2; Length 404;
Best Local Similarity 27.1%; Pred. No. 5.3e-18;
Matches 105; Conservative 52; Mismatches 116; Indels 114; Gaps 18;

QY 16 KVQGFK-----NGPLPQSLQLRLRSGPCHLLLSLGL--GULLLVIIICVGFQNSKFKQ 66
Db 27 QTRGYKSLAGCLGHGPLVLQ-----LLSFTLLAGLLVQVSKVPSSISQESQR 73
QY 67 RDLVTLRTDFSNFTNTVA-----EIQALTSQSSLEETIASL--KAEVEGKQE-- 114
Db 74 QDAI-----YONLTQKAAVAGELSEKSLQEIYQELTQKAAVAGELPEKSKLQEIYQELT 128
QY 115 -RQAGVSEL-----QEHTTQKAHLGHCPCPSVCVPVHSEM----- 149
Db 129 RLKAAVAGELPEKSKLQEIYQELTQKAAVAGELPE--KSKQEIYQELTRLKAAVAGELPEKS 187
QY 150 -----LLRV-----QQLVQDLKKLTQCVATLNN-----NA 174
Db 188 KQEIYQELTRLKAAVAGELPEKSKQEIYQELTRLKAAVAGELPEKSKQEIYQELTQK 247
QY 175 STGTC--CPVNVVEHQDSYFWSHSGMSWAEAEKYCOLKNAHLVINSREEQNF--VQK 230
Db 248 AVERLCHPCPWEWTFPGGNCYFMSNSQRNWHDITACEVGAQLVWIKSAEONFLQLS 307
QY 231 YLGSAYTWMGLSD--PEGAWKWDGTDYATGF--QNNKPGQPDQWQHGGLGGEDCAHFHP 287
Db 308 SRSNRFTWMGLSDLNQEGTWQVDSPLPSPFKQYNNRGEPPN-----VGEEDCAEFGS 361
QY 288 DGRWDDVCQRPYHWVCEAGLQTSQE 314
Db 362 NG-WNDDKCNLAKFWICKKSAASCSRD 387
```

RESULT 11

```
JC7595
scavenger receptor with C-type lectin type I - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C;Accession: JC7595
R;Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001
A;Title: Molecular cloning and functional characterization of a human scavenger receptor
A;Reference number: JC7595; MUID:21092718; PMID:11162630
A;Contents: Placenta
A;Accession: JC7595
A;Molecule type: mRNA
A;Residues: 1-742 <NAK>
A;Cross-references: UNIPROT:Q9BYH7; DDBJ:AB038518
C;Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase receptor family, plays a role in recognizing important role in host defense. It forms a dimer and plays a role in recognizing important role in host defense.
```

```
C;Genetics:
A;Gene: srcl-I
```

```
A;Map position: 18p11.32
C;Keywords: coiled coil; glycoprotein; transmembrane protein
F;1-39/Domain: cytosolic (amino-terminus) #status predicted <CYT>
F;16-19/Region: internalization signal YKRF
F;40-56/Domain: transmembrane #status predicted <TM>
F;57-112/Domain: extracellular #status predicted <EXT>
F;113-335/Domain: coiled coil #status predicted <COC>
F;369-384/Region: serine/threonine-rich #status predicted
F;443-589/Domain: collagen-like #status predicted <COL>
F;607-732/Domain: C-type lectin/carbohydrate recognition domain #status predicted <CRD>
```

Query Match		18.0%; Score 313; DB 2; Length 742;
Best Local Similarity		43.9%; Pred. No. 1.9e-17;
Matches		58; Conservative 26; Mismatches 42; Indels 6; Gaps 4;
QY	177 EGTCCPVNVEHODSCYWFESHGMSWAEAEKVCQLKNAHLVNVNREEQNFKYL-GSA 235	
DB	603 EDNCSPPHKNFTDKCYFSVEKEIFEDAKLFCEDKSSHLVFINTRREEQQWTKQMVGRE 662	
QY	236 YTWMLGSDP--EGAKWVDGTDYATGFQNNKQPPDDWQGHGLGGGEDCAHFHPDGRWMD 293	
DB	663 SHWIGLTDSERENENKWLJGT--SPDYKNWKAGOPDNW-GHGGPGEDCAGLIYAGQWMD 719	
QY	294 DVCQRPYHWCE 305	
DB	720 FOCEDVNNFICE 731	
RESULT 12		
IgE Fc receptor II, low-affinity - rat		
N;Alternate names: CD23; lymphocyte IgE receptor		
C;Species: Rattus norvegicus (Norway rat)		
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004		
C;Accession: S34198		
R;Flores-Romo, L.; Shield, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayal		
submitted to the EMBL Data Library, June 1993		
A;Description: Inhibition of an in vivo antigen-specific IgE response by antibodies to C		
A;Reference number: S34198		
A;Accession: S34198		
A;Molecule type: mRNA		
A;Residues: 1-309 <FLO>		
A;Cross-references: UNIPROT:Q63097; EMBL:X73579; NID:G313672; PIDN:CAA51981.1; PID:G3136		
C;Superfamily: IgE receptor II; C-type lectin homology		
C;Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr		
F;1-25/Domain: intracellular #status predicted <INT>		
F;14-22/Region: stop-transfer sequence		
F;24-46/Domain: transmembrane #status predicted <TM>		
F;7-309/Product: soluble IgE-binding factor (29K) #status predicted <EXT>		
F;149-309/Product: soluble IgE-binding factor (25-27K) #status predicted <BFI>		
F;164-283/Domain: C-type lectin homology <LCH>		
F;192-283, 260-274/Disulfide bonds: #status predicted		
Query Match		16.6%; Score 288.5; DB 1; Length 309;
Best Local Similarity		29.0%; Pred. No. 6.5e-16;
Matches		81; Conservative 47; Mismatches 108; Indels 43; Gaps 11;
QY	47 LGLL-----LVIIICVGFQNSKFORDLVTLRTDFSNTSNTVABTQA-----LTSQSS 96	
DB	29 VGLLTVMVWLLALLLWETEKSLQLGDAAIQNALQMSQNLBELQAEQKMKSDSQ 88	
QY	97 LBETIASLAEVGFQKQERQAGVSELQE--HTTQKAHLGHCHPCPSVCPVHSEMLLRVQ 154	
DB	89 LSQNLNELQEDLINVKSQN-----SELQNLTQLQED-----LVNVKSQGLNEKR 133	
QY	155 QLVQDLKKLTTCQVATL-----NNNASTEGTCPCPVNVVHODSCYWFESHGMSWAEAEKVCQ 210	
DB	134 AASDSLKLEQEEVAKLWIELMSKGTACNVCPKDWLHFQKCYFEGESKQMIQAKFTCS 193	
QY	211 LKNAHLVNVNREEQNFKYLGSAVTWMLSD--PEGAKWVDGTDYATGFQNNKQGPQ 268	
DB	194 DLEGLVSIHSQEQDFLMQHINKKESWIGLQDLNMEGEFVWPDGS--PVGYSNNMPGEP 251	
QY	269 DDWQGHGLGGGEDCAHFHPDGRWDDVCORPY--HWVCE 305	
DB	252 NNG-----GGGEDCVMMRGSGQWMDAFC-RSYLDAWVCE 284	
RESULT 13		
LNMSR		
IgE Fc receptor, low-affinity - mouse		
N;Alternate names: Blast-2; CD23; Fc-epsilon-RII; lymphocyte IgE receptor		
C;Species: Mus musculus (house mouse)		

C;Date: 12-Feb-1993 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004		
C;Accession: A43518; A33840		
R;Gollnick, S.O.; Trounstine, M.L.; Yamashita, L.C.; Kehry, M.R.; Moore, K.W.		
J. Immunol. 144, 1974-1982, 1990		
A;Title: Isolation, characterization, and expression of cDNA clones encoding the mouse		
A;Reference number: A43518; MUID:90171598; PMID:2137845		
A;Accession: A43518		
A;Molecule type: mRNA		
A;Residues: 1-331 <GOL>		
A;Cross-references: UNIPROT:P20693; GB:M34163; NID:gl93242; PIDN:AAA37603.1; PID:G309222		
R;Bettler, B.; Hofstetter, H.; Rao, M.; Yokoyama, W.M.; Kilcherr, F.; Conrad, D.H.		
Proc. Natl. Acad. Sci. U.S.A. 86, 7566-7570, 1989		
A;Title: Molecular structure and expression of the murine lymphocyte low-affinity recept		
A;Reference number: A33840; MUID:90017519; PMID:2529542		
A;Accession: A33840		
A;Molecule type: mRNA		
A;Residues: 1-331 <BET>		
A;Cross-references: GB:M99371; NID:gl93245; PIDN:AAA74898.1; PID:gl93246; GB:M27150		
C;Comment: This receptor for the Fc portion of IgE is expressed in various hematopoietic		
f B-cells.		
C;Superfamily: IgE receptor II; C-type lectin homology		
C;Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr		
F;1-25/Domain: intracellular #status predicted <INT>		
F;14-22/Region: stop-transfer sequence		
F;26-46/Domain: transmembrane #status predicted <TM>		
F;47-331/Domain: extracellular #status predicted <EXT>		
F;66-86/Region: 21-residue repeat		
F;87-107/Region: 21-residue repeat		
F;108-128/Region: 21-residue repeat		
F;129-149/Region: 21-residue repeat		
F;186-305/Domain: C-type lectin homology <LCH>		
F;65,114/Binding site: carbohydrate (Asn) (covalent) #status predicted		
Query Match		16.5%; Score 287; DB 1; Length 331;
Best Local Similarity		28.7%; Pred. No. 9.4e-16;
Matches		82; Conservative 44; Mismatches 128; Indels 32; Gaps 10;
QY	42 LLSLGLGLLLVVICVGFQNSKFORDL-----VTLRTDFSNTSNTVABTQALT 91	
DB	31 LLSTAWAGALLALLLWETEKNLKQLGDTAIGNVSHVT--KDLQKQSNQLAKSQV 88	
QY	92 SQGSSLEETIA---SLKAEVGFQKQERQAGVSELQHTTQKAHLG-HCPHCPSCVCPVHS 147	
DB	89 QMSQNLQELQAEQKQKQADSRLSQNLTCLOEDLNAQSQNSKLSQNLRLQDDLNVNKS 148	
QY	148 EMLLRVQQLVQDLKKLTTCQVATL-----NNNASTEGTCPCPVNVVHODSCYWFESHGMSWA 203	
DB	149 LGLNEKRTASDSLEKQLQEEVAKLWIELISKGTACNICPKNWLHFQKCYFEGKSKQMI 208	
QY	204 EAEKYCQLKNAHLVNVNREEQNFKYLGSAVTWMLSD--PEGAKWVDGTDYATGFQ 261	
DB	209 QARFACSDLQGLVLSIHSQEQDFLMQHINKKDSWIGLQDLNMEGEFVWSDGS--PVGYS 266	
QY	262 NWKQGPDDWQGHGLGGGEDCAHFHPDGRWDDVCORPY--HWVCE 305	
DB	267 NNNPGEPNNG-----GGGEDCVMMRGSGQWMDAFC-RSYLDAWVCE 306	
RESULT 14		
LNCHL		
hepatic lectin - chicken		
C;Species: Gallus gallus (chicken)		
C;Date: 08-Oct-1981 #sequence_revision 08-Oct-1981 #text_change 09-Jul-2004		
C;Accession: A03167; A28194; A40427		
R;Drickamer, K.		
J. Biol. Chem. 256, 5827-5839, 1981		
A;Title: Complete amino acid sequence of a membrane receptor for glycoproteins. Sequence		
A;Reference number: A03167; MUID:81215504; PMID:7240175		
A;Accession: A03167		
A;Molecule type: protein		
A;Residues: 1-207 <DR1>		
A;Cross-references: UNIPROT:P02707		
A;Note: some or all of the cysteines are involved in disulfide bonds		

A>Note: residues 24-48 form an uncharged, hydrophobic region that may interact with or d
 R:Mellow, T.E.; Halberg, D.; Drickamer, K.
 J. Biol. Chem. 263, 5468-5473, 1988
 A>Title: Endocytosis of N-acetylglucosamine-containing glycoproteins by rat fibroblasts
 A:Reference number: A28194; MUID:88186849; PMID:3281941
 A:Accession: A28194
 A:Molecule type: mRNA
 A:Residues: 1-207 <HEL>
 A:Cross-references: GB:J03188; NID:9212246; PIDN:AAA48937.1; PID:9212247
 A:Bezouska, K.; Crichtlow, G.V.; Rose, J.M.; Taylor, M.E.; Drickamer, K.
 J. Biol. Chem. 266, 11604-11609, 1991
 A>Title: Evolutionary conservation of intron position in a subfamily of genes encoding c
 A:Reference number: A40427; MUID:91268022; PMID:2050668
 A:Accession: A40427
 A:Molecule type: DNA
 A:Residues: 1-207 <BEZ>
 A:Cross-references: GB:M63225; GB:M63226; GB:M63227; GB:M63228; GB:M63229; GB:M63230; NI
 C:Comment: Hepatic lectin is a membrane receptor protein that recognizes and binds expos
 and endocytosis.
 C:Genetics:
 A:Introns: 15/1; 50/1; 75/1; 125/3; 163/2
 A:Superfamily: hepatic lectin; C-type lectin homology
 C:Keywords: acetylated amino end; glycoprotein; lectin; transmembrane protein
 F:1-23/Domain: intracellular #status predicted <INT>
 F:24-47/Domain: transmembrane #status predicted <TRA>
 F:48-207/Domain: extracellular #status predicted <EXT>
 F:78-201/Domain: C-type lectin homology <LCH>
 F:1/Modified site: acetylated amino end (Met) #status experimental
 F:67/Binding site: carbohydrate (Asn) (covalent) #status experimental
 Query Match 16.2%; Score 281; DB 1; Length 207; \\
 Best Local Similarity 36.7%; Pred. No. 1.6e-15;
 Matches 62; Conservative 26; Mismatches 67; Indels 14; Gaps 4;
 Db 147 SEMLLRVQQLVQDLKLTQVATLNNASPEGTCPPV-----NWVHQDSCYWFSSHGMS 201
 Db 44 SVSLARIAALSSKLS--TLQSEKPNFSSRDSLLFCGQSQRWYFEGRCYFSLRMS 101
 QY 202 WBAEAKYCOLKNAHLVINSREQNFQVQLGSAITWMLSD--PEGAKWVDGTDYATG 259
 Db 102 WHKAAECSEMHSHLIIDSYAKQNFVFRTRNERFWIGLTDENQGEQWQVVDGTDTRSS 161
 QY 260 FQWKVQGPDDWQGHGLGGEDCAHFPDGRWVDVQCRPHWVCAGL 308
 Db 162 FTFWKEGEPNN-----RGFNEDCAHWTSQWQNDVYCTEYVVCBKL 205
 RESULT 15
 LNHRU
 IGE Fc receptor II, low-affinity [validated] - human
 N:Alternate names: Blast-2; CD23; Fc-epsilon-R-II; lymphocyte IGE receptor
 N:Contains: IGE Fc receptor II, splice form a; IGE Fc receptor II, splice form a'; IGE F
 C:Species: Homo sapiens (man)
 C>Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 09-Jul-2004
 C:Accession: A26067; S03279; S39442; S39443; A26164; A26589; A31924; JL0132; S29107
 R:Kikutani, H.; Inui, S.; Sato, R.; Barsumian, E.L.; Owaki, H.; Yamasaki, K.; Kaisho, T.
 Cell 47, 657-665, 1986
 A>Title: Molecular structure of human lymphocyte receptor for immunoglobulin E.
 A:Reference number: A26067; MUID:87051737; PMID:2877743
 A:Accession: A26067
 A:Molecule type: mRNA
 A:Residues: 1-321 <IKI>
 A:Cross-references: UNIPROT:P06734; GB:M14766; NID:9182449; PIDN:AAA52435.1; PID:9182450
 A:Experimental source: EBV-transformed B lymphoblastoid cells RPMI-8866
 R:Suter, U.; Bastos, R.; Hofstetter, H.
 Nucleic Acids Res. 15, 7295-7308, 1987
 A>Title: Molecular structure of the gene and the 5'-flanking region of the human lymphoc
 A:Reference number: S03279; MUID:88015596; PMID:2958779
 A:Accession: S03279
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 157-284 <SUT>
 A:Cross-references: GB:X06049; NID:931316

A>Note: all exon sequences were determined but the complete sequence is not shown
 R:Matsui, M.; Nunez, R.; Sachi, Y.; Lynch, R.G.; Yodoi, J.
 FEBS Lett. 335, 51-56, 1993
 A>Title: Alternative transcripts of the human CD23/Fc-epsilon-R-II. A possible novel mech
 A:Reference number: S39442; MUID:94063078; PMID:8243664
 A:Accession: S39442
 A:Molecule type: DNA
 A:Residues: 1-7, 'D', 47-50 <MAS1>
 A:Experimental source: splice form a'
 A:Accession: S39443
 A:Molecule type: DNA
 A:Residues: 'MNPPSQD', 47-50 <MAS2>
 A:Experimental source: splice form b'
 A:Rudin, C.; Hofstetter, H.; Sarfati, M.; Levy, C.A.; Suter, U.; Alaimo, D.; Kilcherr, I.
 EMBO J. 6, 109-114, 1987
 A>Title: Cloning and expression of the cDNA coding for a human lymphocyte IGE receptor.
 A:Reference number: A26164; MUID:87218454; PMID:3034567
 A:Accession: A26164
 A:Molecule type: mRNA
 A:Residues: 1-288, 'T', 270-321 <LUD>
 A:Cross-references: GB:X04772; NID:934002; PIDN:CAA28465.1; PID:934003
 A>Note: the codon given for 269-Asn (ACC) is inconsistent with the authors' translation
 R:Kuka, K.; Takami, M.; Kim, C.W.; Honjo, T.; Miyoshi, T.; Tagaya, Y.; Kawabe, T.; Yodoi,
 Proc. Natl. Acad. Sci. U.S.A. 84, 819-823, 1987
 A>Title: Human lymphocyte Fc receptor for IGE: sequence homology of its cloned cDNA with
 A:Reference number: A26589; MUID:87118255; PMID:2949326
 A:Accession: A26589
 A:Molecule type: mRNA
 A:Residues: 1-321 <IKU>
 A:Cross-references: GB:M15059; NID:9182447; PIDN:AAA52434.1; PID:9182448
 A>Note: part of this sequence, including the amino end of soluble forms of the protein, v
 R:Yokota, A.; Kikukani, H.; Tanaka, T.; Sato, R.; Barsumian, E.L.; Suenura, M.; Kishimoto,
 Cell 55, 611-618, 1988
 A>Title: Two species of human Fc-epsilon receptor II (Fc-epsilon-R-II/CD23): tissue-specifi
 A:Reference number: A31924; MUID:89028672; PMID:2972386
 A:Accession: A31924
 A:Molecule type: mRNA
 A:Residues: 'MNPPSQ', 8-14 <YOK>
 A:Cross-references: GB:M23562; NID:9182444
 A:Experimental source: splice form IIB
 R:Letellier, M.; Sarfati, M.; Delespesse, G.
 Mol. Immunol. 26, 1105-1112, 1989
 A>Title: Mechanisms of formation of IGE-binding factors (soluble CD23)-I. Fc epsilon R II
 A:Reference number: JL0132; MUID:90220658; PMID:2534424
 A:Accession: JL0132
 A:Molecule type: protein
 A:Residues: 1-321 <LET>
 A:Experimental source: lymphoblastoid B cell line
 R:Rose, K.; Turcatti, G.; Graber, P.; Pochon, S.; Regamey, P.O.; Jansen, K.U.; Magnenat,
 Biochem. J. 286, 819-824, 1992
 A>Title: Partial characterization of natural and recombinant human soluble CD23.
 A:Reference number: S29107; MUID:93038513; PMID:1417742
 A:Accession: S29107
 A:Molecule type: protein
 A:Residues: 152-166;173-179;189-212;230-263;268-306 <ROS>
 R:Padlan, E.A.; Helm, B.A.
 submitted to the Brookhaven Protein Data Bank, June 1993
 A:Reference number: A51791; PDB:1HLI
 A:Contents: annotation; conformation by theoretical model, residues 173-285
 R:Bajorath, J.
 submitted to the Brookhaven Protein Data Bank, November 1995
 A:Reference number: A65963; PDB:1KJE
 A:Contents: annotation; conformation by theoretical model, residues 173-285
 C:Comment: The sequence of the splice form a is shown.
 C:Comment: This receptor for the Fc portion of IGE is expressed in various hematopoietic
 cells.
 C:Comment: Splice form a is expressed constitutively in B-cells; b is expressed in other
 C:Comment: Soluble IGE-binding factors are produced by proteolytic cleavage of IGE Fc re
 C:Genetics:
 A:Gene: GDB:FCER2; FCE2
 A:Cross-references: GDB:118888; OMIM:151445
 A:Map position: 19p13.3-19p13.3

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 12:44:49 ; Search time 143 Seconds
(without alignments)
762.317 Million cell updates/sec

Title: US-10-829-107-4

Perfect score: 1738

Sequence: 1 MRTYENFQYLENKVKVQGF.....QRYHWVCAGLGQTSQESH 316

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1462099 seqs, 344972447 residues

Total number of hits satisfying chosen parameters: 1462099

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1738	100.0	316	9	US-09-862-802-4
2	1738	100.0	316	16	US-10-829-107-4
3	1718.5	98.9	319	15	US-10-138-588-90
4	1550.5	89.2	289	16	US-10-829-107-13
5	1544	88.8	292	16	US-10-829-107-12
6	1427.5	82.1	273	9	US-09-862-802-10
7	1427.5	82.1	273	16	US-10-829-107-10
8	1410.5	81.2	267	15	US-10-138-588-88
9	923.5	53.1	291	9	US-09-862-802-5
10	923.5	53.1	291	16	US-10-829-107-5
11	783	45.1	287	9	US-09-862-802-6
12	783	45.1	287	15	US-10-236-392-10
13	783	45.1	287	16	US-10-829-107-6

14	760	43.7	306	15	US-10-236-392-8	Sequence 8, Appli
15	754.5	43.4	311	15	US-10-236-392-6	Sequence 6, Appli
16	326.5	18.8	550	15	US-10-072-012-513	Sequence 513, App
17	321	18.5	345	15	US-10-262-839-94	Sequence 94, Appl
18	316	18.2	404	10	US-09-891-894-1	Sequence 1, Appli
19	316	18.2	404	14	US-10-151-274-2	Sequence 2, Appli
20	316	18.2	404	14	US-10-184-150-1	Sequence 1, Appli
21	316	18.2	404	15	US-10-369-214-131	Sequence 131, App
22	316	18.2	404	15	US-10-328-997-1	Sequence 1, Appli
23	313	18.0	342	14	US-10-223-085-28	Sequence 28, Appl
24	313	18.0	342	14	US-10-223-084-28	Sequence 28, Appl
25	313	18.0	342	14	US-10-223-088-28	Sequence 28, Appl
26	313	18.0	342	14	US-10-223-090-28	Sequence 28, Appl
27	313	18.0	342	14	US-10-223-087-28	Sequence 28, Appl
28	313	18.0	342	14	US-10-223-083-28	Sequence 28, Appl
29	313	18.0	342	14	US-10-223-089-28	Sequence 28, Appl
30	313	18.0	342	14	US-10-223-081-28	Sequence 28, Appl
31	313	18.0	342	14	US-10-223-082-28	Sequence 28, Appl
32	313	18.0	342	15	US-10-305-654-28	Sequence 28, Appl
33	313	18.0	342	15	US-10-081-056-28	Sequence 28, Appl
34	313	18.0	457	9	US-09-745-763-218	Sequence 218, App
35	313	18.0	542	9	US-09-745-763-219	Sequence 219, App
36	313	18.0	618	14	US-10-203-860-24	Sequence 24, Appl
37	313	18.0	742	14	US-10-308-448-11	Sequence 11, Appl
38	311	17.9	399	15	US-10-451-459-2	Sequence 2, Appli
39	311	17.9	742	14	US-10-203-860-2	Sequence 2, Appli
40	311	17.9	742	15	US-10-341-434-85	Sequence 85, Appl
41	309.5	17.8	546	15	US-10-072-012-158	Sequence 158, App
42	309.5	17.8	548	15	US-10-072-012-512	Sequence 512, App
43	309.5	17.8	549	15	US-10-072-012-160	Sequence 160, App
44	301	17.3	399	10	US-09-891-894-2	Sequence 2, Appli
45	301	17.3	399	14	US-10-184-150-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-862-802-4
; Sequence 4, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCES: SF06958
; CURRENT APPLICATION NUMBER: US/09/862,802
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein
US-09-862-802-4

Query Match	100.0%;	Score	1738;	DB	9;	Length	316;
Best Local Similarity	100.0%;	Pred. No.	3.1e-151;				
Matches	316;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MRTYENFQYLENKVKVQGF	PLQSLQLRLSGPCHLLLSGLGLLLVTCVVG	60			
DB	1	MRTYENFQYLENKVKVQGF	PLQSLQLRLSGPCHLLLSGLGLLLVTCVVG	60			
QY	61	QNSKFORDLVTRTDFSNFTSNVTVAEIQALTSQGSLEETIASLKAEVEGFKQERQAGVS	120				
DB	61	QNSKFORDLVTRTDFSNFTSNVTVAEIQALTSQGSLEETIASLKAEVEGFKQERQAGVS	120				
QY	121	ELOEHTTOKAHLCCHPCSVCPVHSEMLLRVQOLVODLKLTCCVATLNNASTEGTC	180				

Db 121 ELQHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTCCQVATLNNASTEGTC 180
QY 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAYTWG 240
Db 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAYTWG 240
QY 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQRPY 300
Db 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQRPY 300
QY 301 HWVCEAGLGQTSQESH 316
Db 301 HWVCEAGLGQTSQESH 316

RESULT 2

US-10-829-107-4
; Sequence 4, Application US/10829107
; Publication No. US20040192892A1
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/10/829,107
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/862,802A
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid and protein
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (608)..(673)
; OTHER INFORMATION: short form lacks nucleotides 608-673
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (775)..(776)
; OTHER INFORMATION: ASGPRM (Table 2) has sequence insert encoding GEE between nucleot
; OTHER INFORMATION: ides 775-776
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1064)..(1064)
; OTHER INFORMATION: nucleotide 1064 of DCM2s may be A, which would encode Asn rather
; OTHER INFORMATION: than Asp at the residue numbered 270
US-10-829-107-4

Query Match 100.0%; Score 1738; DB 16; Length 316;
Best Local Similarity 100.0%; Pred. No. 3,1e-151;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTRTYENFQYLENKVKVQGFKNGLPLQSLLOLRSGPCHLLLSLGLGLLLVLCVVG 60
Db 1 MTRTYENFQYLENKVKVQGFKNGLPLQSLLOLRSGPCHLLLSLGLGLLLVLCVVG 60
QY 61 QNSKFORDLVLTDFSNFTSNVTAIEIQAITSQSSLEETIASLKAEEVGFQKQAGVS 120
Db 61 QNSKFORDLVLTDFSNFTSNVTAIEIQAITSQSSLEETIASLKAEEVGFQKQAGVS 120
QY 121 ELQHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTCCQVATLNNASTEGTC 180

Db 121 ELQHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTCCQVATLNNASTEGTC 180
QY 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAYTWG 240
Db 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAYTWG 240
QY 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQRPY 300
Db 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQRPY 300
QY 301 HWVCEAGLGQTSQESH 316
Db 301 HWVCEAGLGQTSQESH 316

RESULT 3

US-10-138-588-90
; Sequence 90, Application US/10138588
; Publication No. US20040018594A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: NOVEL ANTIBODIES THAT BIND TO ANTIGENIC POLYPEPTIDES, NUCLEIC ACI
; FILE REFERENCE: 21402-347A
; CURRENT APPLICATION NUMBER: US/10/138,588
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/288,395
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/308,901
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/313,388
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/324,757
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/288,900
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 203
; SEQ ID NO 90
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-588-90

Query Match 98.9%; Score 1718.5; DB 15; Length 319;
Best Local Similarity 98.7%; Pred. No. 1.9e-149;
Matches 315; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 MTRTYENFQYLENKVKVQGFKNGLPLQSLLOLRSGPCHLLLSLGLGLLLVLCVVG 60
Db 1 MTRTYENFQYLENKVKVQGFKNGLPLQSLLOLRSGPCHLLLSLGLGLLLVLCVVG 60
QY 61 QNSKFORDLVLTDFSNFTSNVTAIEIQAITSQSSLEETIASLKAEEVGFQKQAGVS 120
Db 61 QNSKFORDLVLTDFSNFTSNVTAIEIQAITSQSSLEETIASLKAEEVGFQKQAGVS 120
QY 121 ELQHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTCCQVATLNN---ASTE 177
Db 121 ELQHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTCCQVATLNNNGEASTE 180
QY 178 GTCCPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAYT 237
Db 181 GTCCPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAYT 240
QY 238 WMGLSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQ 297
Db 241 WMGLSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQ 300
QY 298 RPYHWWCEAGLGQTSQESH 316
Db 301 RPYHWWCEAGLGQTSQESH 319

```
RESULT 4
US-10-829-107-13
; Sequence 13, Application US/10829107
; Publication No. US20040192892A1
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/10/829,107
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/862,802A
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein (DCMP2 short form)
US-10-829-107-13
Query Match      89.2%; Score 1550.5; DB 16; Length 289;
Best Local Similarity 91.1%; Pred. No. 4.8e-134;
Matches 288; Conservative 1; Mismatches 0; Indels 27; Gaps 1;

Qy 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSQPCCHLLSLGLGLLLVLCVVG 60
Db 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSQPCCHLLSLGLGLLLVLCVVG 60

Qy 61 QNSKFQRLDLVLTDFSNFTSNTVAETQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120
Db 61 QNSKFQRLDLVLTDFSNFTSNTVAETQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120

Qy 121 ELQEHHTQKAHLGCHPCSPVCPVHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGTC 180
Db 118 -----VHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGTC 153

Qy 181 CPVNWVEHQDSCYWFHSHGSMWAEAEKYCOLKNAHLVWVINSREQNFQKYLGSAYTMWG 240
Db 154 CPVNWVEHQDSCYWFHSHGSMWAEAEKYCOLKNAHLVWVINSREQNFQKYLGSAYTMWG 213

Qy 241 LSDPEGAKWKVDGTDYATGFQNKPKGPDPDQWQHGLGGEDCAHFHPDGRWDDVCORPY 300
Db 214 LSDPEGAKWKVDGTDYATGFQNKPKGPDPDQWQHGLGGEDCAHFHPDGRWDDVCORPY 273

Qy 301 HWVCEAGLGQTSQESH 316
Db 274 HWVCEAGLGQTSQESH 289

RESULT 5
US-10-829-107-12
; Sequence 12, Application US/10829107
; Publication No. US20040192892A1
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/10/829,107
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/862,802A
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 12
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein (ASGPRm is a macrophage derived ASGPR)
US-10-829-107-12
Query Match      88.8%; Score 1544; DB 16; Length 292;
Best Local Similarity 90.6%; Pred. No. 1.9e-133;
Matches 289; Conservative 0; Mismatches 0; Indels 30; Gaps 2;

Qy 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSQPCCHLLSLGLGLLLVLCVVG 60
Db 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSQPCCHLLSLGLGLLLVLCVVG 60

Qy 61 QNSKFQRLDLVLTDFSNFTSNTVAETQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120
Db 61 QNSKFQRLDLVLTDFSNFTSNTVAETQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120

Qy 121 ELQEHHTQKAHLGCHPCSPVCPVHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGTC 177
Db 118 -----VHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGTC 153

Qy 178 GTCCPVNWVEHQDSCYWFHSHGSMWAEAEKYCOLKNAHLVWVINSREQNFQKYLGSAYT 237
Db 154 GTCCPVNWVEHQDSCYWFHSHGSMWAEAEKYCOLKNAHLVWVINSREQNFQKYLGSAYT 213

Qy 238 WMGLSDPEGAKWKVDGTDYATGFQNKPKGPDPDQWQHGLGGEDCAHFHPDGRWDDVCQ 297
Db 214 WMGLSDPEGAKWKVDGTDYATGFQNKPKGPDPDQWQHGLGGEDCAHFHPDGRWDDVCQ 273

Qy 298 RPYHWVCEAGLGQTSQESH 316
Db 274 RPYHWVCEAGLGQTSQESH 292

RESULT 6
US-09-862-802-10
; Sequence 10, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein
US-09-862-802-10
Query Match      82.1%; Score 1427.5; DB 9; Length 273;
Best Local Similarity 84.6%; Pred. No. 9.3e-123;
Matches 270; Conservative 0; Mismatches 0; Indels 49; Gaps 3;
```

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQRLRSGPCHLLSLGLGLLLVLIICVVG 60
DB 1 MTRTYENFOYLENKVKVQGFKNGLPLQS-----L LLLVLIICVVG 41
QY 61 QNSKFORDLVTLTDFSNFTSNVTVAIQALTSQSSLEETIASLKAIEVEGFKQERQAGVS 120
DB 42 QNSKFORDLVTLTDFSNFTSNVTVAIQALTSQSSLEETIASLKAIEVEGFKQERQA--- 98
QY 121 ELQEHHTQKAHLGCHPCFSVCVPVHSEMLLRVQQLVQDLKKLTQOVATLNN---ASTE 177
DB 99 -----VHSEMLLRVQQLVQDLKKLTQOVATLNNGBEASTE 134
QY 178 GTCCPVNWEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAYT 237
DB 135 GTCCPVNWEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAYT 194
QY 238 WMGLSDPEGAWKWVDGTDYATGFQNWKPQDDQWQHGGLGGEDCAHFHPDGRWDDVCQ 297
DB 195 WMGLSDPEGAWKWVDGTDYATGFQNWKPQDDQWQHGGLGGEDCAHFHPDGRWDDVCQ 254
QY 298 RPYHWVCEAGLGQTSQESH 316
DB 255 RPYHWVCEAGLGQTSQESH 273
RESULT 7
US-10-829-107-10
; Sequence 10, Application US/10829107
; Publication No. US20040192892A1
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/10/829,107
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/862,802A
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid and protein
US-10-829-107-10
Query Match 82.1%; Score 1427.5; DB 16; Length 273;
Best Local Similarity 84.6%; Pred. No. 9.3e-123;
Matches 270; Conservative 0; Mismatches 0; Indels 49; Gaps 3;
QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQRLRSGPCHLLSLGLGLLLVLIICVVG 60
DB 1 MTRTYENFOYLENKVKVQGFKNGLPLQS-----L LLLVLIICVVG 41
QY 61 QNSKFORDLVTLTDFSNFTSNVTVAIQALTSQSSLEETIASLKAIEVEGFKQERQAGVS 120
DB 42 QNSKFORDLVTLTDFSNFTSNVTVAIQALTSQSSLEETIASLKAIEVEGFKQERQA--- 98
QY 121 ELQEHHTQKAHLGCHPCFSVCVPVHSEMLLRVQQLVQDLKKLTQOVATLNN---ASTE 177
DB 99 -----VHSEMLLRVQQLVQDLKKLTQOVATLNNGBEASTE 134
QY 178 GTCCPVNWEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAYT 237

DB 135 GTCCPVNWEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAYT 194
QY 238 WMGLSDPEGAWKWVDGTDYATGFQNWKPQDDQWQHGGLGGEDCAHFHPDGRWDDVCQ 297
DB 195 WMGLSDPEGAWKWVDGTDYATGFQNWKPQDDQWQHGGLGGEDCAHFHPDGRWDDVCQ 254
QY 298 RPYHWVCEAGLGQTSQESH 316
DB 255 RPYHWVCEAGLGQTSQESH 273
RESULT 8
US-10-138-588-88
; Sequence 88, Application US/10138588
; Publication No. US20040018594A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: NOVEL ANTIBODIES THAT BIND TO ANTIGENIC POLYPEPTIDES, NUCLEIC ACI
; FILE REFERENCE: 21402-347A
; CURRENT APPLICATION NUMBER: US/10/138,588
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/288,395
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/308,901
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/313,388
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/324,757
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/288,900
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 203
; SEQ ID NO 88
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-588-88
Query Match 81.2%; Score 1410.5; DB 15; Length 267;
Best Local Similarity 89.5%; Pred. No. 3.3e-121;
Matches 263; Conservative 0; Mismatches 4; Indels 27; Gaps 1;
QY 23 GPLPLQSLQRLRSGPCHLLSLGLGLLLVLIICVGFQNSKFORDLVTLTDFSNFTSN 82
DB 1 GPLPLQSLQRLRSGPCHLLSLGLGLLLVLIICVGFQNSKFORDLVTLTDFSNFTSN 60
QY 83 TVAEIQALTSQSSLEETIASLKAIEVEGFKQERQAGVSELOEHTTQKAHLGCHPCFSVC 142
DB 61 TVAEIQALTSQSSLEETIASLKAIEVEGFKQERQA----- 95
QY 143 VPVHSEMLLRVQQLVQDLKKLTQOVATLNNASTEGCCPVNWEHODSCYWFSGMSW 202
DB 96 --VHSEMLLRVQQLVQDLKKLTQOVATLNNASTEGCCPVNWEHODSCYWFSGMSW 153
QY 203 AEAEKYCOLKNAHLVWINSREEQNFVKYLGSAYTWMLGLSDPEGAWKWVDGTDYATGFQ 262
DB 154 AEAEKYCOLKNAHLVWINSREEQNFVKYLGSAYTWMLGLSDPEGAWKWVDGTDYATGFQ 213
QY 263 WKPGQDDQWQHGGLGGEDCAHFHPDGRWDDVCQRPYHWVCEAGLGQTSQESH 316
DB 214 WKPGQDDQWQHGGLGGEDCAHFHPDGRWDDVCQRPYHWVCEAGLGQTSQESH 267
RESULT 9
US-09-862-802-5
; Sequence 5, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B

; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein
US-09-862-802-5

Query Match 53.1%; Score 923.5; DB 9; Length 291;
Best Local Similarity 54.6%; Pred. No. 2.4e-76;
Matches 172; Conservative 47; Mismatches 67; Indels 29; Gaps 5;

QY 1 MTRTYENFOYLENKVK-VQGFKNGLPLOSLLQRLSRGCPCHLLSLGLGLLLVLCVWG 59
DB 1 MTKEYQDLQHLNDESDHQLRGKPPPPQPLQRLCSGPRLLLSLGLSLLLLVVVCVIG 60

QY 60 FQNSKFQRLDLVLTDFSNFTSNVVAEIQALTSQGSLEETIASLKAEEVGFQKQRAQV 119
DB 61 SQNSQLQEBELRGRLRETFNSFTASTAQVKGSLTQGGNVGKMKLSLSQLE--KQK---- 114

QY 120 SELQEHHTQKAHLGCHPCPSVCPVHSEMMLRVQDLKKLTQVATLNNASTEGT 179
DB 115 -DLSED-----HSSLLLVHKQFVSDRLSLSCMAALQNGS--ERT 152

QY 180 CCPNVNVEHQDSYWFSSHGSMWAEAEKYCOLKNAHLVWINSREEQNFVQKYLGSAYTWM 239
DB 153 CCPNVNVEHRSYWFSSRGKAWADADNYCRLEDAHLVWVTSWEEQKFVQHIGPVTW 212

QY 240 GLSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGEDCAHFPHDGRWDDVCORP 299
DB 213 GLHDQNGPWKVDGTDYETGFKNWRPEQPDWYGHGLGGEDCAHFTDDGRWDDVCORP 272

QY 300 YHWVCAGLGQTSQE 314
DB 273 YRWVCETELDKASQE 287

RESULT 10
US-10-829-107-5
; Sequence 5, Application US/10829107
; Publication No. US20040192892A1
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/10/829,107
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/862,802A
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein
US-10-829-107-5

Query Match 53.1%; Score 923.5; DB 16; Length 291;
Best Local Similarity 54.6%; Pred. No. 2.4e-76;
Matches 172; Conservative 47; Mismatches 67; Indels 29; Gaps 5;

QY 1 MTRTYENFOYLENKVK-VQGFKNGLPLOSLLQRLSRGCPCHLLSLGLGLLLVLCVWG 59
DB 1 MTKEYQDLQHLNDESDHQLRGKPPPPQPLQRLCSGPRLLLSLGLSLLLLVVVCVIG 60

QY 60 FQNSKFQRLDLVLTDFSNFTSNVVAEIQALTSQGSLEETIASLKAEEVGFQKQRAQV 119
DB 61 SQNSQLQEBELRGRLRETFNSFTASTAQVKGSLTQGGNVGKMKLSLSQLE--KQK---- 114

QY 120 SELQEHHTQKAHLGCHPCPSVCPVHSEMMLRVQDLKKLTQVATLNNASTEGT 179
DB 115 -DLSED-----HSSLLLVHKQFVSDRLSLSCMAALQNGS--ERT 152

QY 180 CCPNVNVEHQDSYWFSSHGSMWAEAEKYCOLKNAHLVWINSREEQNFVQKYLGSAYTWM 239
DB 153 CCPNVNVEHRSYWFSSRGKAWADADNYCRLEDAHLVWVTSWEEQKFVQHIGPVTW 212

QY 240 GLSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGEDCAHFPHDGRWDDVCORP 299
DB 213 GLHDQNGPWKVDGTDYETGFKNWRPEQPDWYGHGLGGEDCAHFTDDGRWDDVCORP 272

QY 300 YHWVCAGLGQTSQE 314
DB 273 YRWVCETELDKASQE 287

RESULT 11
US-09-862-802-6
; Sequence 6, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein
US-09-862-802-6

Query Match 45.1%; Score 783; DB 9; Length 287;
Best Local Similarity 46.6%; Pred. No. 2e-63;
Matches 146; Conservative 50; Mismatches 89; Indels 28; Gaps 3;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLSRGCPCHLLSLGLGLLLVLCVWG 60
DB 1 MAXDFQDIQQLSSEENDHHPHQGPPPAQPLAQLRCSMVCPSLLALSFNILLVVICVGS 60

QY 61 QNSKFQRLDLVLTDFSNFTSNVVAEIQALTSQGSLEETIASLKAEEVGFQKQRAQV 120
DB 61 QSAQLQAEALRSKAEAFNSFSSLTTEVQAISTHGGSVGDKITSGLAKLEKQQQDLKA--- 117

QY 121 ELQEHHTQKAHLGCHPCPSVCPVHSEMMLRVQDLKKLTQVATLNNASTEGT 180
DB 118 ---DHDLALFLKHP-----VDLRFVACQMEHLHSGS--ORTC 152

QY 181 CPNVNVEHQDSYWFSSHGSMWAEAEKYCOLKNAHLVWINSREEQNFVQKYLGSAYTWM 240
DB 153 CPNVNVEHQDSYWFSSHGSMWAEAEKYCOLKNAHLVWINSREEQNFVQHTNPNFTWIG 212

QY 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGEDCAHFPHDGRWDDVCORP 300

Db 61 QSAQLQAEALSLKEARFNSFSSSTLTEVQAISTHGGSGVDKITSILGAKLEKQOQDLKA --- 117
QY 121 ELQHTTQKAHLGCHPCPVCVPHSEMLLRVQQLVQDLKKLTCQVATLNNASTEGTC 180
Db 118 ---DHDALLFLHKFP-----VDLRFVACQWELLHSGS-QRTC 152
QY 181 CPVNVVEHODSCYWFSGHSGMWAEAEKYCOLKNAHLVWVINSREEQNFVQKYLGSAYTWG 240
Db 153 CPVNVVEHODSCYWFSGHSGMWAEAEKYCOLKNAHLVWVINSREEQNFVQKYLGSAYTWG 212
QY 241 LSDPEGAKWVDGTDYATGQKPKQPDQWQGHGGLGGGDCAHFHPDGRWDDVCORPY 300
Db 213 LTSDSGSKWVGDGTDYRHNKNAVTPDNWGHGHELGSSEDCVEVQPDGRWDDFCLOVY 272
QY 301 HWVCEAGLGQTSQ 313
Db 273 RWVCEKRRNATGE 285

RESULT 14

US-10-236-392-8
; Sequence 8, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladamir
; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRocheville, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard A
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-442A
; CURRENT FILING DATE: 2002-09-06
; CURRENT APPLICATION NUMBER: US/10/236,392
; PRIOR FILING DATE: 2002-03-30
; PRIOR APPLICATION NUMBER: US09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US60/390,155
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US09/635,949
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US60/318,765
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/357,303
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US60/367,753
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US60/369,479
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US09/659,634

; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: US60/318,120
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US60/318,130
; PRIOR FILING DATE: 2001-09-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 794
; SOFTWARE: Custom
; SEQ ID NO 8
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-392-8
Query Match 43.7%; Score 760; DB 15; Length 306;
Best Local Similarity 48.3%; Pred. No. 2.8e-61;
Matches 142; Conservative 44; Mismatches 80; Indels 28; Gaps 3;
QY 20 FKNGPLQLQLRLSGPCHLLSLGLGLLLVILCVGFGNSKFORDLVTLRTDPSNF 79
Db 39 FLKGPPLPAQLRLCSMVCFSLLALSNFILLVILVICVTGSSQSAQLQAEALSLKEARFNSF 98
QY 80 TSNTVAEIQALTSGSSLEBTIASLKAEEVGFQERQAGVSEIQEHTTQKAHLGCHPCP 139
Db 99 SSTLTVEQAIETHGGSGVDKITSILGAKLEKQOQDLKA-----DHDALLFLHKFP--- 149
QY 140 SVCVPVHSEMLLRVQQLVQDLKKLTCQVATLNNASTEGTCPCPVNVVEHODSCYWFSGHSG 199
Db 150 -----VDLRFVACQWELLHSGS-QRTCCPVNVVEHODSCYWFSGHSG 190
QY 200 MSWAEAEKYCOLKNAHLVWVINSREEQNFVQKYLGSAYTWGLSDPEGAKWVDGTDYATG 259
Db 191 KAWAEAEKYCLLENHLVWVINSWEEQKFIQHTNPFNTWGLTSDSGSKWVGDGTDYRHN 250
QY 260 FQNVKPGQPDQWQGHGGLGGGDCAHFHPDGRWDDVCORPYHWVCEAGLGQTSQ 313
Db 251 YKNWAVTPDNWGHGHELGSSEDCVEVQPDGRWDDFCLOVYRWVCEKRRNATGE 304

RESULT 15

US-10-236-392-6
; Sequence 6, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladamir
; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRocheville, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard A

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 12:44:45 ; Search time 22 Seconds
(without alignments)
1072.232 Million cell updates/sec

Title: US-10-829-107-4
Perfect score: 1738
Sequence: 1 MRYTFYFQLENKVKVQGF.....QRPYHWVCZAGLQTSQESH 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1738	100.0	316	3	US-09-111-470-4
2	1738	100.0	316	4	US-09-862-802A-4
3	1550.5	89.2	289	4	US-09-862-802A-13
4	1544	88.8	292	2	US-08-688-342-4
5	1544	88.8	292	2	US-09-113-788-4
6	1544	88.8	292	4	US-09-862-802A-12
7	1427.5	82.1	273	3	US-09-111-470-10
8	1427.5	82.1	273	4	US-09-862-802A-10
9	923.5	53.1	291	2	US-08-688-342-5
10	923.5	53.1	291	2	US-09-113-788-5
11	923.5	53.1	291	3	US-09-111-470-5
12	923.5	53.1	291	4	US-09-862-802A-5
13	923.5	53.1	311	4	US-09-949-016-9962
14	783	45.1	287	3	US-09-111-470-6
15	783	45.1	287	4	US-09-862-802A-6
16	745.5	42.9	272	4	US-09-949-016-9961
17	520	29.9	114	6	US14582-10
18	520	29.9	114	6	US14582-10
19	488	28.1	129	3	US-08-722-126A-10
20	488	28.1	129	5	PCT-US95-04258-10
21	454	26.1	114	6	US14582-8
22	454	26.1	114	6	US14582-8
23	432	24.9	114	6	US14582-14
24	432	24.9	114	6	US14582-14
25	410	23.6	114	6	US14582-13
26	410	23.6	114	6	US14582-13
27	347	20.0	77	4	US-09-531-056A-21

28 316 18.2 404 3 US-09-517-605-2 Sequence 2, Appli
29 287 16.5 327 1 US-08-365-103B-2 Sequence 2, Appli
30 283 16.3 287 1 US-08-365-103B-4 Sequence 4, Appli
31 283 16.3 300 1 US-08-365-103B-6 Sequence 6, Appli
32 279 16.1 292 4 US-09-535-521-2 Sequence 2, Appli
33 279 16.1 292 4 US-09-535-521-5 Sequence 5, Appli
34 278 16.0 208 4 US-09-535-521-20 Sequence 20, Appli
35 276.5 15.9 115 6 5514582-18 Patent No. 5514582
36 276.5 15.9 115 6 5514582-18 Patent No. 5514582
37 274.5 15.8 320 1 US-08-365-103B-10 Sequence 10, Appli
38 274.5 15.7 321 1 US-08-365-103B-8 Sequence 8, Appli
39 273 15.7 219 4 US-09-907-794A-377 Sequence 377, App
40 273 15.7 219 4 US-09-905-125A-377 Sequence 377, App
41 273 15.7 219 4 US-09-902-775A-377 Sequence 377, App
42 273 15.7 219 4 US-09-906-700-377 Sequence 377, App
43 273 15.7 219 4 US-09-903-603A-377 Sequence 377, App
44 273 15.7 219 4 US-09-904-920A-377 Sequence 377, App
45 273 15.7 219 4 US-09-909-064-377 Sequence 377, App

ALIGNMENTS

RESULT 1
US-09-111-470-4
; Sequence 4, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecq, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-111-470-4
Query Match 100.0%; Score 1738; DB 3; Length 316;
Best Local Similarity 100.0%; Pred. No. 1e-167;

```
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLRSGPCHLLLSGLGLLLVLCVVG 60
DB 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLRSGPCHLLLSGLGLLLVLCVVG 60
QY 61 QNSKFQDLVLTLDTSNFTSNFTVAEIQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120
DB 61 QNSKFQDLVLTLDTSNFTSNFTVAEIQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120
QY 121 ELOEHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGTC 180
DB 121 ELOEHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGTC 180
QY 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREONFVKYLGSAYTWG 240
DB 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREONFVKYLGSAYTWG 240
QY 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQRPY 300
DB 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQRPY 300
QY 301 HWVCEAGLGQTSQESH 316
DB 301 HWVCEAGLGQTSQESH 316

RESULT 2
US-09-862-802A-4
; Sequence 4, Application US/09862802A
; Patent No. 6756478
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/09/862,802A
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid and protein
; NAME/KEY: misc feature
; LOCATION: (608)..(673)
; OTHER INFORMATION: short form lacks nucleotides 608-673
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (775)..(776)
; OTHER INFORMATION: ASGPRM (Table 2) has sequence insert encoding GEE between nucleot
; OTHER INFORMATION: ides 775-776
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1064)..(1064)
; OTHER INFORMATION: nucleotide 1064 of DCMF2s may be A, which would encode Asn rather
; OTHER INFORMATION: than Asp at the residue numbered 270
US-09-862-802A-4
Query Match 100.0%; Score 1738; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 1e-167;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLRSGPCHLLLSGLGLLLVLCVVG 60
DB 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLRSGPCHLLLSGLGLLLVLCVVG 60
QY 61 QNSKFQDLVLTLDTSNFTSNFTVAEIQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120
DB 61 QNSKFQDLVLTLDTSNFTSNFTVAEIQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120
QY 121 ELOEHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGTC 180
DB 121 ELOEHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGTC 180
QY 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREONFVKYLGSAYTWG 240
DB 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREONFVKYLGSAYTWG 240
QY 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQRPY 300
DB 241 LSDPEGAKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQRPY 300
QY 301 HWVCEAGLGQTSQESH 316
DB 301 HWVCEAGLGQTSQESH 316

RESULT 3
US-09-862-802A-13
; Sequence 13, Application US/09862802A
; Patent No. 6756478
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/09/862,802A
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 13
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein (DCMP2 short form)
US-09-862-802A-13
Query Match 89.2%; Score 1550.5; DB 4; Length 289;
Best Local Similarity 91.1%; Pred. No. 9.4e-149;
Matches 288; Conservative 1; Mismatches 0; Indels 27; Gaps 1;
QY 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLRSGPCHLLLSGLGLLLVLCVVG 60
DB 1 MTRTYENFOYLENKVKVQGFKNGLPLOSLLQRLRSGPCHLLLSGLGLLLVLCVVG 60
QY 61 QNSKFQDLVLTLDTSNFTSNFTVAEIQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 120
DB 61 QNSKFQDLVLTLDTSNFTSNFTVAEIQALTSQSSLEETIASLKAEEVGFQKQERQAGVS 117
QY 121 ELOEHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGTC 180
DB 118 -----VHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGTC 153
QY 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREONFVKYLGSAYTWG 240
DB 181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREONFVKYLGSAYTWG 240
```

Db 154 CPVNWVEHQDSCYFWSHSGMSWAEAEKYCOLKNAHLVWVINSREQNFVQKYLGSAYT 213
QY 241 LSPDEGAKWVDGTDYATGFGQNNKPGQPDWQGHGLGGEDCAHFHPDGRWDDVCQ 300
Db 214 LSPDEGAKWVDGTDYATGFGQNNKPGQPDWQGHGLGGEDCAHFHPDGRWDDVCQ 273
QY 301 HWCEAGLGQTSQESH 316
Db 274 HWCEAGLGQTSQESH 289

RESULT 4

US-08-688-342-4
; Sequence 4, Application US/08688342
; Patent No. 5871964
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,342
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1235724

US-08-688-342-4
Query Match 88.8%; Score 1544; DB 2; Length 292;
Best Local Similarity 90.6%; Pred. No. 4.4e-148;
Matches 289; Conservative 0; Mismatches 0; Indels 30; Gaps 2;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSGPCHLLLSGLGLLLLVICVWGF 60
Db 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSGPCHLLLSGLGLLLLVICVWGF 60
QY 61 QNSKFQDLVLTTRDTSNFTSNVVAETQALTSQSSLEETIASLKAEVEGFKQERQAGVS 120
Db 61 QNSKFQDLVLTTRDTSNFTSNVVAETQALTSQSSLEETIASLKAEVEGFKQERQAGVS 117
QY 121 ELOEHTTQKAHLGHCPCPSVCPVHSEMMLRRVQDLKKLTCQVATLNNN---ASTE 177
Db 118 -----VHSEMMLRRVQDLKKLTCQVATLNNNGEEASTE 153
QY 178 GTCCPWNVVEHQDSCYFWSHSGMSWAEAEKYCOLKNAHLVWVINSREQNFVQKYLGSAYT 237

Db 154 GTCCPWNVVEHQDSCYFWSHSGMSWAEAEKYCOLKNAHLVWVINSREQNFVQKYLGSAYT 213
QY 238 WMGLSDPEGAKWVDGTDYATGFGQNNKPGQPDWQGHGLGGEDCAHFHPDGRWDDVCQ 297
Db 214 WMGLSDPEGAKWVDGTDYATGFGQNNKPGQPDWQGHGLGGEDCAHFHPDGRWDDVCQ 273
QY 298 RPYHWVCEAGLGQTSQESH 316
Db 274 RPYHWVCEAGLGQTSQESH 292

RESULT 5

US-09-113-788-4
; Sequence 4, Application US/09113788
; Patent No. 5969104
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,788
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/688,342
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1235724

US-09-113-788-4
Query Match 88.8%; Score 1544; DB 2; Length 292;
Best Local Similarity 90.6%; Pred. No. 4.4e-148;
Matches 289; Conservative 0; Mismatches 0; Indels 30; Gaps 2;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSGPCHLLLSGLGLLLLVICVWGF 60
Db 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSGPCHLLLSGLGLLLLVICVWGF 60
QY 61 QNSKFQDLVLTTRDTSNFTSNVVAETQALTSQSSLEETIASLKAEVEGFKQERQAGVS 120
Db 61 QNSKFQDLVLTTRDTSNFTSNVVAETQALTSQSSLEETIASLKAEVEGFKQERQAGVS 117
QY 121 ELOEHTTQKAHLGHCPCPSVCPVHSEMMLRRVQDLKKLTCQVATLNNN---ASTE 177

db 118 -----VHSEMLLRVQQLVQDLKKLTQVATLNNNGBEASTE 153
QY 178 GTCCPVNWEHODSCYWFHSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAVT 237
Db 154 GTCCPVNWEHODSCYWFHSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAVT 213
QY 238 WMGLSDPEGAKWVDGTDYATGFQNWKPQPDWQHGGLGGGEDCAHFPHDPGRWDDVQC 297
Db 214 WMGLSDPEGAKWVDGTDYATGFQNWKPQPDWQHGGLGGGEDCAHFPHDPGRWDDVQC 273
QY 298 RPYHWVCEAGLGQTSQESH 316
Db 274 RPYHWVCEAGLGQTSQESH 292

RESULT 6
US-09-862-802A-12
; Sequence 12, Application US/09862802A
; Patent No. 6756478
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/09/862,802A
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 12
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein (ASGPRm is a macrophage derived ASGPR)
US-09-862-802A-12

Query Match 88.8%; Score 1544; DB 4; Length 292;
Best Local Similarity 90.6%; Pred. No. 4.4e-148;
Matches 289; Conservative 0; Mismatches 0; Indels 30; Gaps 2;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQRLRSRGPCHLLSLGLGLLLVLCVVG 60
Db 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQRLRSRGPCHLLSLGLGLLLVLCVVG 60
QY 61 QNSKFORDLVLTDFSNFTSNVTVAEIQALTSQSSLEETIASLKAEEVGFQKQRAQGS 120
Db 61 QNSKFORDLVLTDFSNFTSNVTVAEIQALTSQSSLEETIASLKAEEVGFQKQRAQ 117
QY 121 ELQEHHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTQVATLNNN 177
Db 118 -----VHSEMLLRVQQLVQDLKKLTQVATLNNNGBEASTE 153
QY 178 GTCCPVNWEHODSCYWFHSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAVT 237
Db 154 GTCCPVNWEHODSCYWFHSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAVT 213
QY 238 WMGLSDPEGAKWVDGTDYATGFQNWKPQPDWQHGGLGGGEDCAHFPHDPGRWDDVQC 297
Db 214 WMGLSDPEGAKWVDGTDYATGFQNWKPQPDWQHGGLGGGEDCAHFPHDPGRWDDVQC 273
QY 298 RPYHWVCEAGLGQTSQESH 316
Db 274 RPYHWVCEAGLGQTSQESH 292

RESULT 7
US-09-111-470-10
; Sequence 10, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)498-1200
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-111-470-10

Query Match 82.1%; Score 1427.5; DB 3; Length 273;
Best Local Similarity 84.6%; Pred. No. 2.6e-136;
Matches 270; Conservative 0; Mismatches 0; Indels 49; Gaps 3;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQRLRSRGPCHLLSLGLGLLLVLCVVG 60
Db 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQRLRSRGPCHLLSLGLGLLLVLCVVG 41
QY 61 QNSKFORDLVLTDFSNFTSNVTVAEIQALTSQSSLEETIASLKAEEVGFQKQRAQGS 120
Db 42 QNSKFORDLVLTDFSNFTSNVTVAEIQALTSQSSLEETIASLKAEEVGFQKQRA 98
QY 121 ELQEHHTTQKAHLGHCHPCPSVCVPVHSEMLLRVQQLVQDLKKLTQVATLNNN 177
Db 99 -----VHSEMLLRVQQLVQDLKKLTQVATLNNNGBEASTE 134
QY 178 GTCCPVNWEHODSCYWFHSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAVT 237
Db 135 GTCCPVNWEHODSCYWFHSGMSWAEAEKYCOLKNAHLVWVINSREEQNFVKYLGSAVT 194
QY 238 WMGLSDPEGAKWVDGTDYATGFQNWKPQPDWQHGGLGGGEDCAHFPHDPGRWDDVQC 297
Db 195 WMGLSDPEGAKWVDGTDYATGFQNWKPQPDWQHGGLGGGEDCAHFPHDPGRWDDVQC 254

QY 298 RPYHWVCEAGLGQTSQESH 316
Db 255 RPYHWVCEAGLGQTSQESH 273

RESULT 8

US-09-862-802A-10
; Sequence 10, Application US/09862802A
; Patent No. 6758478
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.B.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/09/862,802A
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid and protein
US-09-862-802A-10

Query Match 82.1%; Score 1427.5; DB 4; Length 273;
Best Local Similarity 84.6%; Pred. No. 2.6e-136;
Matches 270; Conservative 0; Mismatches 0; Indels 49; Gaps 3;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLSGPCHLLSLGLGLLLVLCVVG 60

Db 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLSGPCHLLSLGLGLLLVLCVVG 41

QY 61 QNSKFORDLTLTDFSNFTNTVAETQALTSOGSSLEETIASLKAEVEGFKQERQV 120

Db 42 QNSKFORDLTLTDFSNFTNTVAETQALTSOGSSLEETIASLKAEVEGFKQERQV 98

QY 121 ELQEHHTQKAHLGHCPCSPVCPVHSEMLLRVQQLVQDLKLTCCOVATLNNN 177

Db 99 -----VHSEMLLRVQQLVQDLKLTCCOVATLNNN 134

QY 178 GTCCPVNVVHQDSCYWFSGMSWAEAEKYCOLKNAHLVIVINSREEQNFVQKYLGSAYT 237

Db 135 GTCCPVNVVHQDSCYWFSGMSWAEAEKYCOLKNAHLVIVINSREEQNFVQKYLGSAYT 194

QY 238 WMGLSDPEGAKWVDGTDYATGFQNKWPGQDDWQGHGLGGGDCAHFHPDGRWDDVCQ 297

Db 195 WMGLSDPEGAKWVDGTDYATGFQNKWPGQDDWQGHGLGGGDCAHFHPDGRWDDVCQ 254

QY 298 RPYHWVCEAGLGQTSQESH 316

Db 255 RPYHWVCEAGLGQTSQESH 273

RESULT 9

US-08-688-342-5
; Sequence 5, Application US/08688342
; Patent No. 5871964
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN

; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,342
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 179079
US-08-688-342-5

Query Match 53.1%; Score 923.5; DB 2; Length 291;
Best Local Similarity 54.6%; Pred. No. 3.5e-85;
Matches 172; Conservative 47; Mismatches 67; Indels 29; Gaps 5;

QY 1 MTRTYENFOYLENKVK-VQGFKNGLPLQSLQLRLSGPCHLLSLGLGLLLVLCVVG 59

Db 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLSGPCHLLSLGLGLLLVLCVVG 60

QY 60 FQNSKFORDLTLTDFSNFTNTVAETQALTSOGSSLEETIASLKAEVEGFKQERQV 119

Db 61 SQNSQLQEEELRLGRLTFSTASTAQVGLSTQGGNVGRKMKLSQLE--KQK--- 114

QY 120 SELQEHHTQKAHLGHCPCSPVCPVHSEMLLRVQQLVQDLKLTCCOVATLNNNASTEGT 179

Db 115 -DUSED-----HSSLLHLVKQFVSDRLSLSCOMALQNGS-ERT 152

QY 180 CCPNVNVVHQDSCYWFSGMSWAEAEKYCOLKNAHLVIVINSREEQNFVQKYLGSAYTWM 239

Db 153 CCPNVNVVHERSCYWFSGMSWAEAEKYCOLKNAHLVIVINSREEQNFVQKYLGSAYTWM 212

QY 240 GLSDPEGAKWVDGTDYATGFQNKWPGQDDWQGHGLGGGDCAHFHPDGRWDDVCQ 299

Db 213 GLHDQNGPKWVDGTDYATGFQNKWPEQDDMYGHGLGGGDCAHFHTDGRWDDVCQ 272

QY 300 YHWVCEAGLGQTSQESH 314

Db 273 YRWVCEAGLGQTSQESH 287

RESULT 10

US-09-113-788-5
; Sequence 5, Application US/09113788
; Patent No. 5969104
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.

/ TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: US
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/113,788
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/688,342
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 291 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 179079
/ US-09-113-788-5

Query Match 53.1%; Score 923.5; DB 2; Length 291;
Best Local Similarity 54.6%; Pred. No. 3.5e-85;
Matches 172; Conservative 47; Mismatches 67; Indels 29; Gaps 5;

QY 1 MTRTYENFOYLENKVK-VQGFKNGLPLQSLQLRLSGPCHLLSLGLGLLLVVCVIG 59
Db 1 MTKEYQDLQHLNDESDHQLRKGPFPPLQLRLCSGPRLLLSGLSLLLVVVCVIG 60

QY 60 FQNSKFORDLVLTLDPSNFTSNVVAEIQAITSQSSLEETIASLKAEEVGFQKQRAQV 119
Db 61 SQNSQLQEBLRLGLRETFNSFTASTAQVKGSLTQGGNVGRKMKLSQLE--KQK- 114

QY 120 SELQHTTQKAHLGCHPCPSVCPVHSEMLLRVQQLVQDLKLTCTOVATLNNNASTGT 179
Db 115 -DLSED-----HSSLHLHVQKFVSDLSLSCMAALQNGS-ERT 152

QY 180 CCPVNVVHSDSCYFWSHSGMSWAEAKYCOLKNAHLVVINSREEQNFVQKYLGSAYTWM 239
Db 153 CCPVNVVHERSCYFWSRSGKAWADADNYCRLEDAHLVVVTSWEEQKFVQHHIGPVNTWM 212

QY 240 GLSDPEGAKWVDGTDYATGFQNKWPGQDDWQHGGLGGEDCAHPHDPGRWDDVCQRP 299
Db 213 GLHDQNGPKWVDGTDYETGFKNWRPEQDDWYGHGLGGEDCAHFTDDGRWDDVCQRP 272

QY 300 YHWVCEAGLGOTSOE 314
Db 273 YRWVCETELDKASQE 287

RESULT 11
US-09-111-470-5
/ Sequence 5, Application US/09111470
/ Patent No. 6277959
/ GENERAL INFORMATION:

/ APPLICANT: Valladeau, Jenny
/ APPLICANT: Ravel, Odile
/ APPLICANT: Bates, Elizabeth E.M.
/ APPLICANT: Ford, John
/ APPLICANT: Saeland, Sem
/ APPLICANT: Lebecque, Serge J.E.
/ TITLE OF INVENTION: Mammalian Membrane Protein Genes;
/ TITLE OF INVENTION: Related Reagents
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DNAX Research Institute
/ STREET: 901 California Avenue
/ CITY: Palo Alto
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94304-1104
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/111,470
/ FILING DATE: 08-JUL-1998
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/053,080
/ FILING DATE: 09-JUL-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ching, Edwin P.
/ REGISTRATION NUMBER: 34,090
/ REFERENCE/DOCKET NUMBER: SF0695
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650)852-9196
/ TELEFAX: (650)496-1200
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 291 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-09-111-470-5

Query Match 53.1%; Score 923.5; DB 3; Length 291;
Best Local Similarity 54.6%; Pred. No. 3.5e-85;
Matches 172; Conservative 47; Mismatches 67; Indels 29; Gaps 5;

QY 1 MTRTYENFOYLENKVK-VQGFKNGLPLQSLQLRLSGPCHLLSLGLGLLLVVCVIG 59
Db 1 MTKEYQDLQHLNDESDHQLRKGPFPPLQLRLCSGPRLLLSGLSLLLVVVCVIG 60

QY 60 FQNSKFORDLVLTLDPSNFTSNVVAEIQAITSQSSLEETIASLKAEEVGFQKQRAQV 119
Db 61 SQNSQLQEBLRLGLRETFNSFTASTAQVKGSLTQGGNVGRKMKLSQLE--KQK- 114

QY 120 SELQHTTQKAHLGCHPCPSVCPVHSEMLLRVQQLVQDLKLTCTOVATLNNNASTGT 179
Db 115 -DLSED-----HSSLHLHVQKFVSDLSLSCMAALQNGS-ERT 152

QY 180 CCPVNVVHSDSCYFWSHSGMSWAEAKYCOLKNAHLVVINSREEQNFVQKYLGSAYTWM 239
Db 153 CCPVNVVHERSCYFWSRSGKAWADADNYCRLEDAHLVVVTSWEEQKFVQHHIGPVNTWM 212

QY 240 GLSDPEGAKWVDGTDYATGFQNKWPGQDDWQHGGLGGEDCAHPHDPGRWDDVCQRP 299
Db 213 GLHDQNGPKWVDGTDYETGFKNWRPEQDDWYGHGLGGEDCAHFTDDGRWDDVCQRP 272

QY 300 YHWVCEAGLGOTSOE 314
Db 273 YRWVCETELDKASQE 287

```
RESULT 12
US-09-862-802A-5
; Sequence 5, Application US/09862802A
; Patent No. 6756478
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth Ester Mary
; APPLICANT: Ford, John
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Saeland, Sem
; TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
; FILE REFERENCE: SF0695 B
; CURRENT APPLICATION NUMBER: US/09/862, 802A
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/053,080
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein
US-09-862-802A-5

Query Match 53.1%; Score 923.5; DB 4; Length 291;
Best Local Similarity 54.6%; Pred. No. 3.5e-85;
Matches 172; Conservative 47; Mismatches 67; Indels 29; Gaps 5;

QY 1 MTRTYENFOYLENKVK-VQGFKNGLPQLQSLRLRSGPCCHLLSLGLGLLLVLCVVG 59
Db 1 MTKEYQDLQHLNDESDHHQLRKGPQPQLQLRCLSGPRLLLSLGLSLLLVVVCVIG 60
QY 60 FQNSKFQDRLVLTDFSNFTNTVAEIOALTQSGSSLEETIASLKAEEVGFQKQERQAGV 119
Db 61 SQNSQLQEELRLGRLRETFNSTASTAQVKGSLTQGGNVGRKMSLSQLE--KQK- 114
QY 120 SELQEHHTQKAHLGCHPCPSVCPVHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGT 179
Db 115 -DLSED-----HSSLLLVKQFVSDRLSLSCMAALQNGS-ERT 152
QY 180 CCPVNVVHODSCYFWSHSGMSWAEAEKYCOLKNAHLVINSREEQNFVKYLGSAITWM 239
Db 153 CCPVNVVHERSCYFWSRSGKAWADADNYCRLEDAHLVVVTSWEEQKFVQHHIGPVNTWM 212
QY 240 GLSDPEGAKWVDGTDYATGQFNWKPGQDDWQGHGLGGEDCAHFHPDGRWDDVCORP 299
Db 213 GLHDQGPWKWVDGTDYETGFKNWRPEQPDWYGHGLGGEDCAHFTDGRWDDVCORP 272
QY 300 YHWVCEAGLGQTSQE 314
Db 273 YRWVCETELDKASQE 287

RESULT 13
US-09-949-016-9962
; Sequence 9962, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9962
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9962

Query Match 53.1%; Score 923.5; DB 4; Length 311;
Best Local Similarity 54.6%; Pred. No. 3.9e-85;
Matches 172; Conservative 47; Mismatches 67; Indels 29; Gaps 5;

QY 1 MTRTYENFOYLENKVK-VQGFKNGLPQLQSLRLRSGPCCHLLSLGLGLLLVLCVVG 59
Db 21 MTKEYQDLQHLNDESDHHQLRKGPQPQLQLRCLSGPRLLLSLGLSLLLVVVCVIG 80
QY 60 FQNSKFQDRLVLTDFSNFTNTVAEIOALTQSGSSLEETIASLKAEEVGFQKQERQAGV 119
Db 81 SQNSQLQEELRLGRLRETFNSTASTAQVKGSLTQGGNVGRKMSLSQLE--KQK- 134
QY 120 SELQEHHTQKAHLGCHPCPSVCPVHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGT 179
Db 135 -DLSED-----HSSLLLVKQFVSDRLSLSCMAALQNGS-ERT 172
QY 180 CCPVNVVHODSCYFWSHSGMSWAEAEKYCOLKNAHLVINSREEQNFVKYLGSAITWM 239
Db 173 CCPVNVVHERSCYFWSRSGKAWADADNYCRLEDAHLVVVTSWEEQKFVQHHIGPVNTWM 232
QY 240 GLSDPEGAKWVDGTDYATGQFNWKPGQDDWQGHGLGGEDCAHFHPDGRWDDVCORP 299
Db 233 GLHDQGPWKWVDGTDYETGFKNWRPEQPDWYGHGLGGEDCAHFTDGRWDDVCORP 292
QY 300 YHWVCEAGLGQTSQE 314
Db 293 YRWVCETELDKASQE 307

RESULT 14
US-09-111-470-6
; Sequence 6, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
```

NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0695
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 287 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-111-470-6

Query Match 45.1%; Score 783; DB 3; Length 287;
Best Local Similarity 46.6%; Pred. No. 6.1e-71;
Matches 146; Conservative 50; Mismatches 89; Indels 28; Gaps 3;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSGPCHLLLSLGLGLLLVLCVVG 60
DB 1 MAKDFODIQQLSSEENDHPFHQPPAPQPLAQLCLCSMVCFSLLALSFNILLVVICVTGS 60
QY 61 QNSKFORDLVTLRTDPSNFTSNTVAEIQALTSOGSSLEETIASLKAEBEGFKQERQAGVS 120
DB 61 QSAQLQAEIQLSLKEAFSNFSSSTLVEVQAIETHGSGVGDKITSLGAKLEKQQQDLKA--- 117
QY 121 ELQEHHTQKAHLGHCHPCSVCPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGTC 180
DB 118 ---DHDALLFHLKHFP-----VDLRFVACQOMELLHSGS-QRTC 152
QY 181 CPVNWVEHQDSCYWFSGMSWAEAEKYCOLKNAHLVWINSREBONFVKYLGSAITWNG 240
DB 153 CPVNWVEHQGSCYWFSGKAWAEAEKYCOLKNAHLVWINSWEEQKFIVQHTNPNFTWIG 212
QY 241 LSDPEGANKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCORPY 300
DB 213 LTDSGSKWVDGTDYRHNKNAWVTPDNWHGHELGSSEDCVEVQPDGRWDDFCLOVY 272
QY 301 HWVCEAGLGQTSQ 313
DB 273 RWVCEKRRNATGE 285

Search completed: May 27, 2005, 12:49:19
Job time : 24 secs

NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0695
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 287 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-111-470-6

Query Match 45.1%; Score 783; DB 3; Length 287;
Best Local Similarity 46.6%; Pred. No. 6.1e-71;
Matches 146; Conservative 50; Mismatches 89; Indels 28; Gaps 3;

QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQLRLRSGPCHLLLSLGLGLLLVLCVVG 60
DB 1 MAKDFODIQQLSSEENDHPFHQPPAPQPLAQLCLCSMVCFSLLALSFNILLVVICVTGS 60
QY 61 QNSKFORDLVTLRTDPSNFTSNTVAEIQALTSOGSSLEETIASLKAEBEGFKQERQAGVS 120
DB 61 QSAQLQAEIQLSLKEAFSNFSSSTLVEVQAIETHGSGVGDKITSLGAKLEKQQQDLKA--- 117
QY 121 ELQEHHTQKAHLGHCHPCSVCPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGTC 180
DB 118 ---DHDALLFHLKHFP-----VDLRFVACQOMELLHSGS-QRTC 152
QY 181 CPVNWVEHQDSCYWFSGMSWAEAEKYCOLKNAHLVWINSREBONFVKYLGSAITWNG 240
DB 153 CPVNWVEHQGSCYWFSGKAWAEAEKYCOLKNAHLVWINSWEEQKFIVQHTNPNFTWIG 212
QY 241 LSDPEGANKWVDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCORPY 300
DB 213 LTDSGSKWVDGTDYRHNKNAWVTPDNWHGHELGSSEDCVEVQPDGRWDDFCLOVY 272
QY 301 HWVCEAGLGQTSQ 313
DB 273 RWVCEKRRNATGE 285

RESULT 15
US-09-862-802A-6
Sequence 6, Application US/09862802A
Patent No. 6756478
GENERAL INFORMATION:
APPLICANT: Valladeau, Jenny
APPLICANT: Ravel, Odile
APPLICANT: Bates, Elizabeth Ester Mary
APPLICANT: Ford, John
APPLICANT: Lebecque, Serge J.E.
APPLICANT: Saeland, Sem
TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes; Related Reagents
FILE REFERENCE: SF0695 B
CURRENT APPLICATION NUMBER: US/09/862, 802A
CURRENT FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/053,080
PRIOR FILING DATE: 1997-07-09
PRIOR APPLICATION NUMBER: US 09/111,470
PRIOR FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 287
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: mammalian protein
US-09-862-802A-6

XX Dendritic cell membrane proteins - used to treat conditions associated
PT with abnormal physiology or development.
XX
XX
XX Claim 2; Page 68-69; 82pp; English.
XX
XX Dendritic cell membrane protein 1 (DCMP1) and DCMP2 nucleic acids can be
CC used as markers for distinguishing cell types, including genomic aspects
CC of cells, as well as mRNA and protein expression patterns. They can also
CC be used to detect chromosomal abnormalities. The proteins can be used to
CC diagnose disorders associated with expression misregulation. They can
CC also be used to treat conditions associated with abnormal physiology or
CC development, including abnormal proliferation, e.g. cancerous conditions
CC or degenerative conditions. Abnormal proliferation, regeneration,
CC degeneration and atrophy may be modulated using the proteins. The
CC proteins may also play a role in regulation or development of
CC haematopoietic cells
XX

SQ	Sequence 316 AA;	
	Query Match 100.0%; Score 1738; DB 2; Length 316;	
	Best Local Similarity 100.0%; Pred. No. 1.9e-154;	
	Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MTRTYENFOYLENKVKVQGFKNQPLPLQSLQRLRSGPCHLLLSLGLGILLVLCVVG 60	
DB	1 MTRTYENFOYLENKVKVQGFKNQPLPLQSLQRLRSGPCHLLLSLGLGILLVLCVVG 60	
QY	61 QNSKFORDLVLTIRDFSNFTSNTVAEIQALTSQSSLEETIASLKAEEVGFQKQAGVS 120	
DB	61 QNSKFORDLVLTIRDFSNFTSNTVAEIQALTSQSSLEETIASLKAEEVGFQKQAGVS 120	
QY	121 ELQEHHTTQKAHLGHCPCPSVCPVHSEMLLRVQDLVKLTQVATLNNASTEGTC 180	
DB	121 ELQEHHTTQKAHLGHCPCPSVCPVHSEMLLRVQDLVKLTQVATLNNASTEGTC 180	
QY	181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAVTW 240	
DB	181 CPVNWVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAVTW 240	
QY	241 LSDPEGAWKWVDGTDYATGFQNWKPQDDWQHGGLGGEDCAHFHPDGRWDDVCQ 300	
DB	241 LSDPEGAWKWVDGTDYATGFQNWKPQDDWQHGGLGGEDCAHFHPDGRWDDVCQ 300	
QY	301 HWCEAGLGQTSQESH 316	
DB	301 HWCEAGLGQTSQESH 316	
RESULT 2		
ID	ABG66703 standard; protein; 319 AA.	
XX	ABG66703;	
AC	ABG66703;	
DT	30-AUG-2002 (first entry)	
XX	Human novel polypeptide #38.	
DE		
XX	Human; inflammatory condition; shock; sepsis; immune response; cancer;	
KW	wound healing; central nervous system disease; haematopoesis;	
KW	peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;	
KW	myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;	
KW	cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;	
KW	bone degenerative disorder; periodontal disease; reperfusion injury;	
KW	lung fibrosis; liver fibrosis; autoimmune disease; bacterial infection;	
KW	allergic condition; thrombolysis; thrombosis; coagulation disorder;	
XX	fungal infection.	
OS	Homo sapiens.	
XX		
PN	WO200244340-A2.	
XX		
PD	06-JUN-2002.	
XX		
PF	30-NOV-2001; 2001WO-US047004.	
XX		
PR	30-NOV-2000; 2000US-00728952.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;	
PI	Yamazaki V, Ujwal ML, Drmanac RT;	
XX		
DR	WPI; 2002-508509/54.	
DR	N-PSDB; ABK94927.	
XX		
XX	Novel nucleic acids and polypeptides for diagnosis, treatment of	
PT	inflammatory, autoimmune, nervous system, myeloid or lymphoid cell	
PT	disorders, cancer and promoting wound healing.	
XX		

PS	Claim 10; Page 607-608; 672pp; English.	
XX		
CC	The invention relates to human novel polynucleotides and associated	
CC	polypeptides. The polynucleotides and polypeptides are useful for	
CC	treating inflammatory conditions such as arthritis, nephritis, Crohn's	
CC	disease, ischaemia-reperfusion injury, shock, sepsis, immune responses	
CC	and cancer and for promoting wound healing. The sequences are used to	
CC	induce the proliferation of neural cells and regeneration of nerve and	
CC	brain tissue, and are useful for the treatment of central and peripheral	
CC	nervous system diseases and neuropathies, such as Alzheimer's disease,	
CC	Parkinson's disease, Huntington's disease and amyotrophic lateral	
CC	sclerosis. The sequences are involved in chemotactic or chemokinetic	
CC	activity, regulation of haematopoesis, treatment of myeloid or lymphoid	
CC	cell disorders and platelet disorders such as thrombocytopenia,	
CC	regeneration of bone, cartilage, tendon, ligament and/or nerve tissue	
CC	growth, tissue repair, healing of burns, incisions, ulcers, treatment of	
CC	osteoporosis, osteoarthritis, bone degenerative disorders and periodontal	
CC	disease. The sequences of the invention are also useful for gut	
CC	protection or regeneration and treatment of lung or liver fibrosis,	
CC	reperfusion injury in various tissues, immune deficiencies and disorders	
CC	including severe combined immunodeficiency (SCID), bacterial or fungal	
CC	infections, autoimmune disorders e.g. multiple sclerosis and myasthenia	
CC	gravis, allergic conditions such as asthma, thrombolysis or thrombosis	
CC	and coagulation disorders. Sequences ABG66666-ABG66758 represent human	
CC	novel polypeptides of the invention	
XX		
SQ	Sequence 319 AA;	
	Query Match 98.9%; Score 1718.5; DB 5; Length 319;	
	Best Local Similarity 98.7%; Pred. No. 1.3e-152;	
	Matches 315; Conservative 0; Mismatches 1; Indels 3; Gaps 1;	
QY	1 MTRTYENFOYLENKVKVQGFKNQPLPLQSLQRLRSGPCHLLLSLGLGILLVLCVVG 60	
DB	1 MTRTYENFOYLENKVKVQGFKNQPLPLQSLQRLRSGPCHLLLSLGLGILLVLCVVG 60	
QY	61 QNSKFORDLVLTIRDFSNFTSNTVAEIQALTSQSSLEETIASLKAEEVGFQKQAGVS 120	
DB	61 QNSKFORDLVLTIRDFSNFTSNTVAEIQALTSQSSLEETIASLKAEEVGFQKQAGVS 120	
QY	121 ELQEHHTTQKAHLGHCPCPSVCPVHSEMLLRVQDLVKLTQVATLNNASTEGTC 177	
DB	121 ELQEHHTTQKAHLGHCPCPSVCPVHSEMLLRVQDLVKLTQVATLNNASTEGTC 180	
QY	178 GTCCPVNVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAVT 237	
DB	181 GTCCPVNVEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAVT 240	
QY	238 WMLSDPEGAWKWVDGTDYATGFQNWKPQDDWQHGGLGGEDCAHFHPDGRWDDVCQ 297	
DB	241 WMLSDPEGAWKWVDGTDYATGFQNWKPQDDWQHGGLGGEDCAHFHPDGRWDDVCQ 300	
QY	298 RPYHWVCEAGLGQTSQESH 316	
DB	301 RPYHWVCEAGLGQTSQESH 319	
RESULT 3		
ID	ABU69170 standard; protein; 319 AA.	
XX	ABU69170;	
AC	ABU69170;	
DT	02-JUN-2003 (first entry)	
XX	Human NOVX polypeptide #45.	
DE		
XX	Human; NOVX; metabolic disorder; diabetes; infectious disease; obesity;	
KW	anorexia; cancer; cardiovascular disorder; asthma; neurogenesis;	
KW	neurodegenerative disorder; epilepsy; immune disorder; osteoarthritis;	
KW	haematopoietic disorder; inflammatory skin disorder; dyslipidemia;	
KW	haematopoesis; wound healing; angiogenesis; bacterial infection;	
KW	viral infection; fungal infection; helminthic infection; atherosclerosis;	

KW protozoal infection; hypertension.
 XX Homo sapiens.
 OS
 XX
 PN WQ200290504-A2.
 XX
 XX
 PD 14-NOV-2002.
 XX
 XX
 PF 02-MAY-2002; 2002WO-US014342.
 XX
 XX 03-MAY-2001; 2001US-0288395P.
 PR 04-MAY-2001; 2001US-0288900P.
 PR 07-MAY-2001; 2001US-0289087P.
 PR 14-MAY-2001; 2001US-0290753P.
 PR 15-MAY-2001; 2001US-0291189P.
 PR 16-MAY-2001; 2001US-0291243P.
 PR 18-MAY-2001; 2001US-0292001P.
 PR 21-MAY-2001; 2001US-0292374P.
 PR 22-MAY-2001; 2001US-0292587P.
 PR 23-MAY-2001; 2001US-0293107P.
 PR 29-MAY-2001; 2001US-0294110P.
 PR 30-MAY-2001; 2001US-0294434P.
 PR 31-MAY-2001; 2001US-0294827P.
 PR 18-JUN-2001; 2001US-0298988P.
 PR 31-JUL-2001; 2001US-0308901P.
 PR 17-AUG-2001; 2001US-0313388P.
 PR 21-AUG-2001; 2001US-0313851P.
 PR 17-SEP-2001; 2001US-0313937P.
 PR 17-SEP-2001; 2001US-0322701P.
 PR 17-SEP-2001; 2001US-0322802P.
 PR 25-SEP-2001; 2001US-0324757P.
 PR 27-SEP-2001; 2001US-0325314P.
 PR 27-SEP-2001; 2001US-0325682P.
 PR 21-NOV-2001; 2001US-0332129P.
 PR 03-DEC-2001; 2001US-0336882P.
 PR 14-DEC-2001; 2001US-0340305P.
 PR 01-MAY-2002; 2002US-00138588.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
 PI Chapoval A, Edinger S, Gerlach V, Gorman L, Gunther E, Guo X;
 PI Kekuda R, Lepley DM, Li L, Liu X, Malyankar UM, Miller CE;
 PI Millet I, Padigaru M, Patturajan M, Pena CBA, Rieger DK, Shenoy SG;
 PI Shinkets RA, Spytek KA, Taupier RJ, Vernet CM, Voss EZ;
 PI Zerhusen BD;
 XX
 DR WPI; 2003-103512/09.
 DR N-PSDB; ACA10155.
 XX
 PT New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX
 PS Claim 2; Page 170; 340pp; English.
 XX
 CC The invention relates to human NOVX polypeptides and the polynucleotides
 CC encoding them. The polypeptides, polynucleotides and antibodies that bind
 CC immunospecifically to the polypeptides are useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease,
 CC preferably a NOVX-associated disorder. The sequences are useful for
 CC treating, preventing or diagnosing diseases such as metabolic disorders,
 CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,
 CC helminthic, and protozoal), anorexia, cancer, cardiovascular disorders
 CC (e.g. hypertension, atherosclerosis), neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease), epilepsy, immune disorders,
 CC osteoarthritis, haematopoietic disorders, inflammatory skin disorders,
 CC asthma and various dyslipidemias. The nucleic acids and polypeptides may
 CC also be used as targets for the identification of small molecules that
 CC modulate or inhibit e.g. neurogenesis, cell differentiation, cell
 CC proliferation, haematopoiesis, wound healing and angiogenesis, and in the
 CC generation of antibodies that bind immunospecifically to NOVX substances

CC for use in therapeutic or diagnostic methods. The nucleic acids are
 CC further used as hybridisation probes, and in chromosome mapping, tissue
 CC typing, preventive medicine and pharmacogenomics. Sequences ABU69126-
 CC ABU69171 represent human NOVX polypeptides of the invention
 XX
 SQ Sequence 319 AA;
 Query Match 98.9%; Score 1718.5; DB 6; Length 319;
 Best Local Similarity 98.7%; Pred. No. 1.3e-152;
 Matches 315; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 QY 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQRLRSGPCHLLSLGLGLLLVLCVVG 60
 DB 1 MTRTYENFOYLENKVKVQGFKNGLPLQSLQRLRSGPCHLLSLGLGLLLVLCVVG 60
 QY 61 QNSKFQDLVTLRTDFSNFTSNVTVAIEIQALTSQSSLEETIASLKAEVEGFKQERQAGVS 120
 DB 61 QNSKFQDLVTLRTDFSNFTSNVTVAIEIQALTSQSSLEETIASLKAEVEGFKQERQAGVS 120
 QY 121 ELQEHHTQKAHLGCHPCPSVCYPVHSEMILLRQVLVQDLKLTCCQVATLNNN---ASTE 177
 DB 121 ELQEHHTQKAHLGCHPCPSVCYPVHSEMILLRQVLVQDLKLTCCQVATLNNN---ASTE 180
 QY 178 GTCCPVNWVEHQDSYWFHSHGMSWAEAEKYCOLKNAHLVIVINSRERQNFVKYLGSAYT 237
 DB 181 GTCCPVNWVEHQDSYWFHSHGMSWAEAEKYCOLKNAHLVIVINSRERQNFVKYLGSAYT 240
 QY 238 WMGLSDPEGAKWVDGTDYATGFQNKWKPQDDMQHGLGGGEDCAHFHPDGRWDDVCQ 297
 DB 241 WMGLSDPEGAKWVDGTDYATGFQNKWKPQDDMQHGLGGGEDCAHFHPDGRWDDVCQ 300
 QY 298 RPYHWVCEAGLGQTSQESH 316
 DB 301 RPYHWVCEAGLGQTSQESH 319
 RESULT 4
 ADO08343
 ID ADO08343 standard; protein; 315 AA.
 XX
 AC ADO08343;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Human NOVX polypeptide #45.
 XX
 KW Human; NOVX; cardiomyopathy; atherosclerosis; hypertension; scleroderma;
 KW obesity; cancer; diabetes; haemophilia; graft-versus-host disease; AIDS;
 KW asthma; Crohn's disease; multiple sclerosis; infection; anorexia;
 KW cancer-associated cachexia; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; haematopoietic disorder;
 KW wasting disorder.
 XX
 OS Homo sapiens.
 XX
 PN US2004018594-A1.
 XX
 PD 29-JAN-2004.
 XX
 PF 01-MAY-2002; 2002US-00138588.
 XX
 PR 03-MAY-2001; 2001US-0288395P.
 PR 04-MAY-2001; 2001US-0288900P.
 PR 07-MAY-2001; 2001US-0289087P.
 PR 14-MAY-2001; 2001US-0290753P.
 PR 15-MAY-2001; 2001US-0291189P.
 PR 16-MAY-2001; 2001US-0291243P.
 PR 18-MAY-2001; 2001US-0292001P.
 PR 21-MAY-2001; 2001US-0292374P.
 PR 22-MAY-2001; 2001US-0292587P.
 PR 23-MAY-2001; 2001US-0293107P.
 PR 29-MAY-2001; 2001US-0294110P.
 PR 30-MAY-2001; 2001US-0294434P.

PR 31-MAY-2001; 2001US-0294827P.
 PR 31-JUL-2001; 2001US-0308901P.
 PR 17-AUG-2001; 2001US-0313388P.
 PR 21-AUG-2001; 2001US-0313851P.
 PR 21-AUG-2001; 2001US-0313937P.
 PR 17-SEP-2001; 2001US-0322701P.
 PR 17-SEP-2001; 2001US-0322802P.
 PR 25-SEP-2001; 2001US-0324757P.
 PR 27-SEP-2001; 2001US-0325314P.
 PR 27-SEP-2001; 2001US-0325682P.
 PR 01-NOV-2001; 2001US-0332129P.
 PR 03-DEC-2001; 2001US-0336882P.
 PR 14-DEC-2001; 2001US-0340305P.
 XX (ALSO/) ALSOBROOK J P.
 PA (ANDE/) ANDERSON D W.
 PA (BOLD/) BOLDOG F L.
 PA (BURG/) BURGESS C E.
 PA (CASW/) CASMAN S J.
 PA (CHAF/) CHAPOVAL A.
 PA (EDIN/) EDINGER S R.
 PA (GERL/) GERLACH V.
 PA (GORM/) GORMAN L.
 PA (GUNT/) GUNTHER E.
 PA (GUOX/) GUO X S.
 PA (KEKU/) KEKODA R.
 PA (LEPL/) LEPLEY D M.
 PA (LILL/) LI L.
 PA (LIUX/) LIU X.
 PA (MALY/) MALYANKAR U M.
 PA (MILL/) MILLER C E.
 PA (MILL/) MILLER I.
 PA (PADI/) PADIGARU M.
 PA (PATT/) PATTURAJAN M.
 PA (PENA/) PENA C E A.
 PA (RIEG/) RIEGER D K.
 PA (SHEN/) SHENOY S G.
 PA (SHIM/) SHIMKETS R A.
 PA (SPYT/) SPYTEK K A.
 PA (TAUP/) TAUPIER R J.
 PA (VERN/) VERNET C A M.
 PA (VOSS/) VOSS E Z.
 PA (ZERH/) ZERHUSEN B D.
 XX
 PI Chabrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
 PI Chapoval A, Edinger SR, Gerlach V, Gorman L, Gunther E, Guo XS;
 PI Kekuda R, Lepley DM, Li L, Liu X, Malyankar UM, Miller CE;
 PI Millet I, Padigaru M, Patturajan M, Pena CE, Rieger DK, Shenoy SG;
 PI Shimkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ;
 PI Zerhusen BD;
 XX
 WPI: 2004-122037/12.
 DR N-PSDB; ADO08342.
 XX
 PT New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, diabetes,
 PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
 PT or pharmacogenomics.
 XX
 PS Claim 2; SEQ ID NO 90; 219pp; English.
 XX
 CC The invention relates to human NOVX polypeptides and the polynucleotides
 CC encoding them. The polypeptides, polynucleotides and antibodies that bind
 CC immunospecifically to the polypeptides are useful in diagnosing, treating
 CC or preventing NOVX-associated disorders such as cardiomyopathy,
 CC atherosclerosis, hypertension, scleroderma, obesity, cancer, diabetes,
 CC haemophilia, graft-versus-host disease, AIDS, asthma, Crohn's disease,
 CC multiple sclerosis, infections, anorexia, cancer-associated cachexia,
 CC neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's
 CC disease), haematopoietic disorders and wasting disorders. The
 CC polynucleotides are also used as hybridisation probes, in chromosome
 CC mapping and in tissue typing. The polypeptides are also useful as
 CC vaccines. This sequence represents a human NOVX polypeptide of the

CC invention.
 XX
 SQ Sequence 315 AA;
 Query Match 95.0%; Score 1651.5; DB 8; Length 315;
 Best Local Similarity 96.9%; Pred. No. 2.6e-146; Mismatches 5; Indels 3;
 Matches 308; Conservative 0;
 QY 1 MTRTYENFOYLENKVKVQGFKNKGNPLPLOSLLQRLRSRSGPCHLLLSLGLGLLLVLCVVG 60
 DB 1 MTRTYENFOYLENKVKVQGFKNKGNPLPLOSLLQRLRSRSGPCHLLLSLGLGLLLVLCVVG 60
 QY 61 QNSKFORDLVLTDFSNFTNTVAEIQALTSQSSLEETIASLKAEVEGFKQERQAGVS 120
 DB 61 QNSKFORDLVLTDFSNFTNTVAEIQALTSQSSLEETIASLK-VEGFKQERQAGVS 118
 QY 121 ELOHHTTQKAHLGHCPCPSVCVPVHSEMLLRVOOLVQDLKLTQVATIAN--NASTEG 178
 DB 119 ELOHHTTQK-HLGHCPCPCPSVCVPVHSEMLLRVOOLVQDLKLTQVATLNNGEASTEG 177
 QY 179 TCCPVNWEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAATW 238
 DB 178 TCCPVNWEHODSCYWFSGMSWAEAEKYCOLKNAHLVWINSREEQNFVKYLGSAATW 237
 QY 239 MGLSDPEGAWKWDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQR 298
 DB 238 MGLSDPEGAWKWDGTDYATGFQNWKPQDDWQGHGLGGGEDCAHFHPDGRWDDVCQR 297
 QY 299 PYHWVCEAGLGQTSQESH 316
 DB 298 PYHQVCEAGKGQTSQESH 315
 RESULT 5
 AAW88129
 ID AAW88129 standard; protein; 273 AA.
 XX
 AC AAW88129;
 XX
 DT 11-MAY-1999 (first entry)
 XX
 DE Variant primate DCM2 C-lectin family gene protein sequence.
 KW Primate; dendritic cell membrane protein; DCMPI; DCM2;
 KW chromosomal abnormality; expression misregulation;
 KW abnormal proliferation; regeneration; degeneration; haematopoietic cell.
 XX
 OS Mammalia.
 XX
 PN WO9902562-A1.
 XX
 PD 21-JAN-1999.
 XX
 PF 08-JUL-1998; 98WO-US013436.
 XX
 PR 09-JUL-1997; 97US-0053080P.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Valladeau J, Ravel O, Bates EEM, Ford J, Saeland S, Lebecque SJE;
 XX
 DR WPI: 1999-120786/10.
 DR N-PSDB; AAX04868.
 XX
 PT Dendritic cell membrane proteins - used to treat conditions associated
 PT with abnormal physiology or development.
 XX
 PS Claim 2(b); Page 76-78; 82pp; English.
 XX
 CC Dendritic cell membrane protein 1 (DCMP1) and DCM2 nucleic acids can be
 CC used as markers for distinguishing cell types, including genomic aspects
 CC of cells, as well as mRNA and protein expression patterns. They can also
 CC be used to detect chromosomal abnormalities. The proteins can be used to

CC diagnose disorders associated with expression misregulation. They can
 CC also be used to treat conditions associated with abnormal physiology or
 CC development, including abnormal proliferation, e.g. cancerous conditions
 CC or degenerative conditions. Abnormal proliferation, regeneration,
 CC degeneration and atrophy may be modulated using the proteins. The
 CC proteins may also play a role in regulation or development of
 CC haematopoietic cells
 XX
 SQ Sequence 273 AA;

Query Match 82.1%; Score 1427.5; DB 2; Length 273;
 Best Local Similarity 84.6%; Pred. No. 2.4e-125;
 Matches 270; Conservative 0; Mismatches 0; Indels 49; Gaps 3;

QY 1 MTRTYENFOYLENKVKVQGFNGPLPLQSLQRLRSRSGCHLLLSGLGLLLVLCVVG 60
 Db 1 MTRTYENFOYLENKVKVQGFNGPLPLQSLQRLRSRSGCHLLLSGLGLLLVLCVVG 41
 QY 61 QNSKFORDLVTLRTDSNFTSNTVAEIQALTSQSSLEETIASLKAEEVGFQKQERQGV 120
 Db 42 QNSKFORDLVTLRTDSNFTSNTVAEIQALTSQSSLEETIASLKAEEVGFQKQERQ 98
 QY 121 ELQETHTTQKAHLGCHPCSPVVPVHSEMLLRVQQLVQDLKCLTCQVATLNNN---ASTE 177
 Db 99 -----VHSEMLLRVQQLVQDLKCLTCQVATLNNNGEASTE 134
 QY 178 GTCCPVNWEHQDSYWFHSGMSWAEAEKYCOLKNAHLVINSRBEQNFVQKYLGSAYT 237
 Db 135 GTCCPVNWEHQDSYWFHSGMSWAEAEKYCOLKNAHLVINSRBEQNFVQKYLGSAYT 194
 QY 238 WMGLSPPEGAKWVDGTDYATGFQNKWPGQDDWQHGGLGGEDCAHFHDPGRWDDVCQ 297
 Db 195 WMGLSPPEGAKWVDGTDYATGFQNKWPGQDDWQHGGLGGEDCAHFHDPGRWDDVCQ 254
 QY 298 RPYHWVCEAGLGQTSQESH 316
 Db 255 RPYHWVCEAGLGQTSQESH 273

RESULT 6
 ABU69169
 ID ABU69169 standard; protein; 267 AA.
 XX
 AC ABU69169;
 XX
 DT 02-JUN-2003 (first entry)
 XX
 DE Human NOVX polypeptide #44.
 XX
 KW Human; NOVX; metabolic disorder; diabetes; infectious disease; obesity;
 KW anorexia; cancer; cardiovascular disorder; asthma; neurogenesis;
 KW neurodegenerative disorder; epilepsy; immune disorder; osteoarthritis;
 KW haematopoietic disorder; inflammatory skin disorder; dyslipidemia;
 KW haematopoiesis; wound healing; angiogenesis; bacterial infection;
 KW viral infection; fungal infection; helminthic infection; atherosclerosis;
 KW protozoal infection; hypertension.
 XX
 OS Homo sapiens.
 XX
 PN W0200290504-A2.
 XX
 PD 14-NOV-2002.
 XX
 XX 02-MAY-2002; 2002WO-US014342.
 XX
 XX 03-MAY-2001; 2001US-0288395P.
 PR 04-MAY-2001; 2001US-0288900P.
 PR 07-MAY-2001; 2001US-0289087P.
 PR 14-MAY-2001; 2001US-0290753P.
 PR 15-MAY-2001; 2001US-0291189P.
 PR 16-MAY-2001; 2001US-0291243P.
 PR 18-MAY-2001; 2001US-0292011P.
 PR 21-MAY-2001; 2001US-0292374P.

PR 22-MAY-2001; 2001US-0292587P.
 PR 23-MAY-2001; 2001US-0293107P.
 PR 29-MAY-2001; 2001US-0294110P.
 PR 30-MAY-2001; 2001US-0294434P.
 PR 31-MAY-2001; 2001US-0294827P.
 PR 18-JUN-2001; 2001US-0298988P.
 PR 31-JUL-2001; 2001US-0308901P.
 PR 17-AUG-2001; 2001US-0313388P.
 PR 21-AUG-2001; 2001US-0313851P.
 PR 21-AUG-2001; 2001US-0313937P.
 PR 17-SEP-2001; 2001US-0322701P.
 PR 17-SEP-2001; 2001US-0322802P.
 PR 25-SEP-2001; 2001US-0324757P.
 PR 27-SEP-2001; 2001US-0325314P.
 PR 27-SEP-2001; 2001US-0325682P.
 PR 21-NOV-2001; 2001US-0332129P.
 PR 03-DEC-2001; 2001US-0336882P.
 PR 14-DEC-2001; 2001US-0340305P.
 PR 01-MAY-2002; 2002US-00138588.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 PI Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
 PI Chapoval A, Edinger S, Gerlach V, Gorman L, Gunther E, Guo X;
 PI Kekuda R, Lepley DM, Li L, Liu X, Malyankar UM, Miller CE;
 PI Millet I, Padigaru M, Patturajan M, Pena CE, Rieger DK, Shenoy SG;
 PI Shimkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ;
 PI Zerhusen BD;
 XX
 DR WPI; 2003-103512/09.
 DR N-PSDB; ACA10154.
 XX
 PT New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX
 PS Claim 2; Page 169; 340pp; English.
 XX
 CC The invention relates to human NOVX polypeptides and the polynucleotides
 CC encoding them. The polypeptides, polynucleotides and antibodies that bind
 CC immunospecifically to the polypeptides are useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease,
 CC preferably a NOVX-associated disorder. The sequences are useful for
 CC treating, preventing or diagnosing diseases such as metabolic disorders,
 CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,
 CC helminthic, and protozoal), anorexia, cancer, cardiovascular disorders
 CC (e.g. hypertension, atherosclerosis), neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease), epilepsy, immune disorders,
 CC osteoarthritis, haematopoietic disorders, inflammatory skin disorders,
 CC asthma and various dyslipidemias. The nucleic acids and polypeptides may
 CC also be used as targets for the identification of small molecules that
 CC modulate or inhibit e.g. neurogenesis, cell differentiation, cell
 CC proliferation, haematopoiesis, wound healing and angiogenesis, and in the
 CC generation of antibodies that bind immunospecifically to NOVX substances
 CC for use in therapeutic or diagnostic methods. The nucleic acids are
 CC further used as hybridisation probes, and in chromosome mapping, tissue
 CC typing, preventive medicine and pharmacogenomics. Sequences ABU69126-
 CC ABU69171 represent human NOVX polypeptides of the invention
 XX
 SQ Sequence 267 AA;

Query Match 81.2%; Score 1410.5; DB 6; Length 267;
 Best Local Similarity 89.5%; Pred. No. 9.1e-124;
 Matches 263; Conservative 0; Mismatches 4; Indels 27; Gaps 1;

QY 23 GPLPLQSLQRLRSRSGCHLLLSGLGLLLVLCVVGQNSKFORDLVTLRTDSNFTSN 82
 Db 1 GPLPLQSLQRLRSRSGCHLLLSGLGLLLVLCVVGQNSKFORDLVTLRTDSNFTSN 60
 QY 83 TVAEIQALTSQSSLEETIASLKAEEVGFQKQERQGVSELTQKALHGHCHPCSPVC 142
 Db 61 TVAEIQALTSQSSLEETIASLKAEEVGFQKQERQGVSELTQKALHGHCHPCSPVC 95

```
QY 143 VPHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGTCCPVNVVBEHQDSCYWFSGMSW 202
DB 96 --VHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGTCCPVNVVBEHQDSCYWFSGMSW 153
QY 203 ABAEKYCOLKNAHLVVINSREBQNFVQKYLGSAYTWMGLSDPEGAKWVDGTDYATGFQN 262
DB 154 ABAEKYCOLKNAHLVVINSREBQNFVQKYLGSAYTWMGLSDPEGAKWVDGTDYATGFQN 213
QY 263 WKPGQDDQGHGLGGEDCAHFHPDGRWDDVCQRPYHVWCEAGLGQTSQESH 316
DB 214 WKPGQDDQGHGLGGEDCAHFHPVGRWDDVCQRPYHVWCEAGLGQTSQESH 267

RESULT 7
ADO08341
ID ADO08341 standard; protein; 267 AA.
XX
AC ADO08341;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human NOVX polypeptide #44.
XX
KW Human; NOVX; cardiomyopathy; atherosclerosis; hypertension; scleroderma;
KW obesity; cancer; diabetes; haemophilia; graft-versus-host disease; AIDS;
KW asthma; Crohn's disease; multiple sclerosis; infection; anorexia;
KW cancer-associated cachexia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; haematopoietic disorder;
KW wasting disorder.
XX
OS Homo sapiens.
XX
PN US2004018594-A1.
XX
PD 29-JAN-2004.
XX
PF 01-MAY-2002; 2002US-00138588.
XX
PR 03-MAY-2001; 2001US-0288395P.
PR 04-MAY-2001; 2001US-0288900P.
PR 07-MAY-2001; 2001US-0289087P.
PR 14-MAY-2001; 2001US-0290753P.
PR 15-MAY-2001; 2001US-0291189P.
PR 16-MAY-2001; 2001US-0291243P.
PR 18-MAY-2001; 2001US-0292001P.
PR 21-MAY-2001; 2001US-0292374P.
PR 22-MAY-2001; 2001US-0292587P.
PR 23-MAY-2001; 2001US-0293107P.
PR 29-MAY-2001; 2001US-0294110P.
PR 30-MAY-2001; 2001US-0294434P.
PR 31-MAY-2001; 2001US-0294827P.
PR 31-JUL-2001; 2001US-0308901P.
PR 17-AUG-2001; 2001US-031388P.
PR 21-AUG-2001; 2001US-0313851P.
PR 21-AUG-2001; 2001US-0313937P.
PR 17-SEP-2001; 2001US-0322701P.
PR 17-SEP-2001; 2001US-0322802P.
PR 25-SEP-2001; 2001US-0324757P.
PR 27-SEP-2001; 2001US-0325314P.
PR 27-SEP-2001; 2001US-0325682P.
PR 21-NOV-2001; 2001US-0332129P.
PR 03-DEC-2001; 2001US-0336882P.
PR 14-DEC-2001; 2001US-0340305P.
XX
(Also/) ALSOBROOK J P.
PA (ANDE/) ANDERSON D W.
PA (BOLD/) BOLDOGF L.
PA (BURG/) BURGESS C E.
PA (CASM/) CASMAN S J.
PA (CHAP/) CHAFOVAL A.
PA (EDIN/) EDINGER S R.
PA (GERL/) GERLACH V.
```

```
PA (GORM/) GORMAN L.
PA (GUNT/) GUNTHER E.
PA (GUOX/) GUO X S.
PA (KEKU/) KEKUDA R.
PA (LEPL/) LEFLEY D M.
PA (LILL/) LI L.
PA (LILU/) LIU X.
PA (MALY/) MALYANKAR U M.
PA (MILL/) MILLER C E.
PA (MILL/) MILLER I.
PA (PADI/) PADIGARU M.
PA (PATT/) PATTURAJAN M.
PA (PENA/) PENA C E A.
PA (RIEG/) RIEGER D K.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHIMKETS R A.
PA (SPYT/) SPYTEK K A.
PA (TAUP/) TAUPIER R J.
PA (VERN/) VERNET C A M.
PA (VOSS/) VOSS E Z.
PA (ZERH/) ZERHUSEN B D.
XX
PI Alsbrook JP, Anderson DW, Boldog FL, Burgees CE, Casman SJ;
PI Chapoval A, Edinger SR, Gerlach V, Gorman L, Gunther E, Guo XS;
PI Kekuda R, Lefley DM, Li L, Liu X, Malyankar UM, Miller CE;
PI Millet I, Padigar M, Patturajan M, Pena CEA, Rieger DK, Shenoy SG;
PI Shmkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ;
PI Zerhusen BD;
XX
DR WPI; 2004-122037/12.
DR N-PSDB; ADO08340.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes,
PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
PT or pharmacogenomics.
XX
PS Claim 2; SEQ ID NO 88; 219pp; English.
XX
CC The invention relates to human NOVX polypeptides and the polynucleotides
CC encoding them. The polypeptides, polynucleotides and antibodies that bind
CC immunospecifically to the polypeptides are useful in diagnosing, treating
CC or preventing NOVX-associated disorders such as cardiomyopathy,
CC atherosclerosis, hypertension, scleroderma, obesity, cancer, diabetes,
CC haemophilia, graft-versus-host disease, AIDS, asthma, Crohn's disease,
CC multiple sclerosis, infections, anorexia, cancer-associated cachexia,
CC neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's
CC disease), haematopoietic disorders and wasting disorders. The
CC polynucleotides are also used as hybridisation probes, in chromosome
CC mapping and in tissue typing. The polypeptides are also useful as
CC vaccines. This sequence represents a human NOVX polypeptide of the
CC invention.
XX
SQ Sequence 267 AA;
Query Match 81.2%; Score 1410.5; DB 8; Length 267;
Best Local Similarity 89.5%; Pred. No. 9.1e-124;
Matches 263; Conservative 0; Mismatches 4; Indels 27; Gaps 1;
QY 23 GPLPLQSLQLRLRGPCCHLLSLGLGLLLLVICVGFQNSKFORDLVTLRTDTSNFTSN 82
DB 1 GPLPLQSLQLRLRGPCCHLLSLGLGLLLLVICVGFQNSKFORDLVTLRTDTSNFTSN 60
QY 83 TVAETQALTSQGSLEETIASIKAEVGFQKORAGVSELOEHTTQKAHLGHCHPCPSVC 142
DB 61 TVAETQALTSQGSLEETIASIKAEVGFQKORQA----- 95
QY 143 VPHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGTCCPVNVVBEHQDSCYWFSGMSW 202
DB 96 --VHSEMLLRVQQLVQDLKLTCCQVATLNNASTEGTCCPVNVVBEHQDSCYWFSGMSW 153
QY 203 ABAEKYCOLKNAHLVVINSREBQNFVQKYLGSAYTWMGLSDPEGAKWVDGTDYATGFQN 262
```


SQ	Sequence 314 AA;	
	Query Match	51.9%; Score 902; DB 8; Length 314;
	Best Local Similarity	50.9%; Pred. No. 7.4e-76;
	Matches 172; Conservative	47; Mismatches 67; Indels 52; Gaps 6;
QY	1 MTRTYENFOYLENKVK-VQGFKNGLPQLQSLRLSGPCHLLSLGLGLLLVVICVVG 59	
DB	1 MTKEYQDLQHLNDESDHHQLRKGPFPQPLQLRLCSGPRLLLSLGLSLLLVVVCVIG 60	
QY	60 FQNSKFORDLVLTDFSNFTNTVAEIQALTSQSSLEETIASLKAVEGFKQERQAGV 119	
DB	61 SQNSQLOEBELGLRLETFTASTEAQVKGSTQGGNVGRKMKLSLESOLE--KQOK- 114	
QY	120 SELQHTTQKAHLGHCHPCSPVCPVHSEMLLRVQQLVODLKKLTCCOVATLNNASTEGT 179	
DB	115 -DLSED-----HSSLLLVHKQFVSDLSLSCQMAALQGNKS-ERT 152	
QY	180 CCPVNVVHSDSCYWFSGHSGMAEAEKYCOLKNAHLVVINSREE----- 224	
DB	153 CCPVNVVHERSCYWFSGKAWADADNYCRLEDAHLVVVTSWEEOVTRRRVWEAGWPRR 212	
QY	225 -----QNFVKYLGSAIYTWGLSDPEGAKWVDGTDYATGQFNWKPGQDDWQGHGL 276	
DB	213 DHPPSLSKQFVQHHTGIPVNTWGLHDQNGPWKVDGTDYETGFKNWRPEQDDWYGHGL 272	
QY	277 GGEDCAHFHPDGRWDDVCORPYHWHVCEAGLGOTSQE 314	
DB	273 GGEDCAHFTDDGRWDDVCORPYRWVCETELDKASQE 310	
RESULT 13		
ID	ABM82571	
XX	ID ABM82571 standard; protein; 290 AA.	
AC	ABM82571;	
XX	18-NOV-2004 (first entry)	
XX	Human diagnostic and therapeutic pprotein SEQ ID NO:2820.	
XX	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.	
XX	Homo sapiens.	
XX	WO2004023973-A2.	
XX	25-MAR-2004.	
XX	12-SEP-2003; 2003WO-US028227.	
XX	12-SEP-2002; 2002US-0410259P.	
XX	12-SEP-2002; 2002US-0410260P.	
XX	(INCY-) INCYTE CORP.	
XX	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen P;	
XX	Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;	
XX	Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy IP;	
XX	Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;	
XX	Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;	
XX	Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;	
XX	Xu Y, Kwong M, Pollicky JL, Hurwitz BL, Ma Y, Jackson JL, Gletzen D;	
XX	Patury S, Shi X, Suarez CJ;	
XX	WPI; 2004-329368/30.	
XX	N-PSDB; ACN41223.	
XX	New diagnostic and therapeutic polynucleotides and polypeptides, useful	
XX	in diagnosing a condition, disease or disorder associated with human	
XX	molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or	
XX	in gene mapping.	
PS	Claim 27; Page; 190pp; English.	
XX	The invention relates to novel diagnostic and therapeutic polynucleotides	
CC	selected from one of the 2722 sequences defined in the specification. A	
CC	polynucleotide of the invention may have a use in gene therapy. The human	
CC	diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be	
CC	used to diagnose a particular condition, disease or disorder associated	
CC	with human molecules, e.g. cell proliferative disorders,	
CC	autoimmune/inflammatory disorder, developmental disorder, endocrine	
CC	disorder, neurological disorders, gastrointestinal disorders, or	
CC	infections caused by virus, bacteria, fungi or parasite. The dithp	
CC	molecules may also be used in genetic mapping, in identifying individuals	
CC	from minute biological samples, in detecting single nucleotide	
CC	polymorphisms, as molecular weight markers, and for somatic or germ-line	
CC	gene therapy. The present sequence represents a dithp protein of the	
CC	invention. Note: The sequence data for this patent is not represented in	
CC	the printed specification, but was obtained in electronic format directly	
CC	from WIPO at www.wipo.int/pct/en/sequences/listing.htm	
XX	Sequence 290 AA;	
SQ	Query Match	49.6%; Score 862; DB 8; Length 290;
	Best Local Similarity	48.5%; Pred. No. 3.8e-72;
	Matches 164; Conservative	37; Mismatches 61; Indels 76; Gaps 4;
QY	1 MTRTYENFOYLENKVK-VQGFKNGLPQLQSLRLSGPCHLLSLGLGLLLVVICVVG 59	
DB	1 MTKEYQDLQHLNDESDHHQLRKGPFPQPLQLRLCSGPRLLLSLGLSLLLVVVCVIG 60	
QY	60 FQNSKFORDLVLTDFSNFTNTVAEIQALTSQSSLEETIASLKAVEGFKQERQAGV 119	
DB	61 SQNSQLOEBELGLRLETFTASTEAQVKGSTQGGNVGRKMKLSLESOLE----- 95	
QY	120 SELQHTTQKAHLGHCHPCSPVCPVHSEMLLRVQQLVODLKKLTCCOVATLNNASTEGT 179	
DB	96 -----HSSLLLVHKQFVSDLSLSCQMAALQGNKS-ERT 128	
QY	180 CCPVNVVHSDSCYWFSGHSGMAEAEKYCOLKNAHLVVINSREE----- 224	
DB	129 CCPVNVVHERSCYWFSGKAWADADNYCRLEDAHLVVVTSWEEOVTRRRVWEAGWPRR 188	
QY	225 -----QNFVKYLGSAIYTWGLSDPEGAKWVDGTDYATGQFNWKPGQDDWQGHGL 276	
DB	189 DHPPSLSKQFVQHHTGIPVNTWGLHDQNGPWKVDGTDYETGFKNWRPEQDDWYGHGL 248	
QY	277 GGEDCAHFHPDGRWDDVCORPYHWHVCEAGLGOTSQE 314	
DB	249 GGEDCAHFTDDGRWDDVCORPYRWVCETELDKASQE 286	
RESULT 14		
ID	ABM82570	
XX	ID ABM82570 standard; protein; 290 AA.	
XX	ABM82570;	
XX	18-NOV-2004 (first entry)	
XX	Human diagnostic and therapeutic pprotein SEQ ID NO:2819.	
XX	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.	
XX	Homo sapiens.	
XX	WO2004023973-A2.	
XX	25-MAR-2004.	
XX	12-SEP-2003; 2003WO-US028227.	
XX	12-SEP-2002; 2002US-0410259P.	
XX	12-SEP-2002; 2002US-0410260P.	

PA (INCY-) INCYTE CORP.
 XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
 PI Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;
 PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
 PI Paralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patry S, Shi X, Suarez CJ;
 XX WPI: 2004-329368/30.
 DR N-PSDB; ACN41222.
 XX
 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 XX Claim 27; Page: 190pp; English.
 XX
 XX The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 CC
 XX Sequence 290 AA;
 XX
 Query Match 49.6%; Score 862; DB 8; Length 290;
 Best Local Similarity 48.5%; Pred. No. 3.8e-72;
 Matches 164; Conservative 37; Mismatches 61; Indels 76; Gaps 4;
 QY 1 MTRTYENFOYLENKVK-VQGFKNGLPLQSLQLRLSGPCHLLSLGLGLLLVLCVWG 59
 DB 1 MTRKEYQDLQHLNDESDHQLRKGPPLPQLQLRCSGPRLLLSLGLSLLLVWVCVIG 60
 QY 60 FQNSKFORDLVLTIRDTFSNFTSNTVAIEIQALTSQGSLEETIASLKAEVGFQKQAGV 119
 DB 61 SQNSQLQELRLGLRETFNSFTASTEAQVKGSLTGGVGVGRKMSLSQLE--KQOK--- 95
 QY 120 SELQHTTQKAHLGHCPHCPSCVCPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGT 179
 DB 96 -----HSSLLLVKQFVSURLSUSCOMAALQNGS-ERT 128
 QY 180 CCPVNWYEHQDSYWFHSGHGWAEAEKYCOLKNAHLVWINSREE----- 224
 DB 129 CCPVNWYEHRSYWFHSGHGWADADNYCRLEDAHLVWVTSWEEQVTRTRVWEAGWPRR 188
 QY 225 -----QNFVKYLGSAVTWGLSDPEGAKWVDGTDVATGPNQWKPQDPDQGHGL 276
 DB 189 DHPPPSLSQKRFVQHHIGPVNTWMLHDQNGQPKWVDGTDYETGPKNWRPEQPDWYGHGL 248
 QY 277 GGGEDCAHFHPDGRWDDVQCPVHVCYCEAGLQTSOE 314
 DB 249 GGGEDCAHFTDDGRWDDVQCPVHVCYCEAGLQTSOE 286
 RESULT 15
 ID AAW15250 standard; protein; 274 AA.
 XX

AC AAW15250;
 XX 17-OCT-2003 (revised)
 DT 23-NOV-1997 (first entry)
 XX Asialoglycoprotein receptor H1 cytoplasmic+extracellular domains.
 DE Asialoglycoprotein receptor H1; AGPR; autoimmune hepatitis; autoantibody;
 XX diagnosis.
 KW Asialoglycoprotein receptor H1; AGPR; autoimmune hepatitis; autoantibody;
 XX diagnosis.
 OS Homo sapiens.
 OS Chimeric.
 XX
 XX Location/Qualifiers
 FT Domain 1..42
 FT /label= Cytoplasmic_domain
 FT Domain 43..274
 FT /label= Extracellular_domain
 XX
 PN EP773289-A2.
 XX
 XX 14-MAY-1997.
 XX
 XX 20-AUG-1996; 96EP-00113349.
 XX
 XX 21-AUG-1995; 95JP-00212118.
 XX
 XX (TOFU) TONEN CORP.
 XX
 XX Tanida E, Ohue C, Yagi S, Hasegawa A, Kiyosawa K, Yano A;
 XX WPI: 1997-261316/24.
 DR
 XX
 XX Asialo:glyco:protein receptor H1 and L-H2 soluble derivatives - comprise
 PT extracellular domains, optionally also with cytoplasmic domains, useful
 PT for autoimmune hepatitis diagnosis.
 XX
 XX Claim 3; Page: 40pp; English.
 XX
 XX This polypeptide sequence comprises the cytoplasmic (CTD) and
 CC extracellular (ECD) domains of human full-length asialoglycoprotein
 CC receptor (AGPR) H1 (see AAW15245). It can be produced on a large scale,
 CC in a form free of contaminating liver antigens, by PCR amplification (see
 CC AAR6962-65) of CTD+ECD DNA (see also AAR6950), and expression in
 CC bacterial (pref. E. coli) or animal (pref. mammalian) host cells. The ECD
 CC alone (AAW15249) and AGPR L-H2 derivatives (see AAW15251-52) are also
 CC claimed. The appearance of autoantibodies against AGPR can be used as an
 CC indicator for autoimmune hepatitis (AIH), a disease that can lead to
 CC cirrhosis and fatal intractable hepatitis. Recombinant soluble AGPR
 CC derivatives can be used in a claimed method for detecting and measuring
 CC anti-AGPR antibodies, e.g. by ELISA, Western blotting, etc. This allows
 CC AIH to be diagnosed and distinguished from other diseases. AGPR
 CC derivatives or full-length receptors can also be used for the analysis
 CC and purification of asialoglycoproteins and to develop inhibiting agents
 CC against asialoglycoprotein incorporation, or viral and other protein
 CC invasion, into liver cells. (Updated on 17-OCT-2003 to standardise OS
 CC field)
 XX
 SQ Sequence 274 AA;
 XX
 Query Match 48.2%; Score 838; DB 2; Length 274;
 Best Local Similarity 50.2%; Pred. No. 6.3e-70;
 Matches 158; Conservative 46; Mismatches 65; Indels 46; Gaps 6;
 QY 1 MTRTYENFOYLENKVK-VQGFKNGLPLQSLQLRLSGPCHLLSLGLGLLLVLCVWG 59
 DB 1 MTRKEYQDLQHLNDESDHQLRKGPPLPQLQLRCSGPRLLLSLGLSLLLVWVCVIG 43
 QY 60 FQNSKFORDLVLTIRDTFSNFTSNTVAIEIQALTSQGSLEETIASLKAEVGFQKQAGV 119
 DB 44 SQNSQLQELRLGLRETFNSFTASTEAQVKGSLTGGVGVGRKMSLSQLE--KQOK--- 97
 QY 120 SELQHTTQKAHLGHCPHCPSCVCPVHSEMLLRVQQLVQDLKKLTQCVATLNNASTEGT 179

